# INVESTIGATING THE ROLE OF SEX LETHAL IN DROSOPHILA GERMLINE SEXUAL IDENTITY

by

Shekerah C. Primus

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

July, 2016

**Abstract**

Sex determination is an essential process in sexually reproducing species. Defects in sex determination lead to disorders of sex development and infertility. Therefore, studying sex determination in different species is key to improve our understanding of this important process. In *Drosophila melanogaster*, the way in which germ cells determine their sexual identity is different from and less understood than somatic sex determination. To improve our understanding of *Drosophila* germline sex determination, I investigated the molecular function of *Sex lethal* (*Sxl*), the key gene promoting female identity in the germline. Specifically, I performed an RNA expression profiling experiment (RNA-Seq) to identify targets of SXL in the germline. This RNA-Seq revealed a few hundred genes with significant changes in expression upon loss of *Sxl* from the germline. This data also showed that *Sxl* plays a largely repressive role in the germline, as a larger number of genes became upregulated in response to loss of *Sxl*. One such gene is *PHD Finger Protein 7*, which was previously shown to promote male identity in the germline.

Additionally a previously uncharacterized gene, *CG15930*, was validated as a target of SXL in the germline. Specifically this gene is highly expressed in males and strongly repressed by SXL in females. *CG15930* mutant males have reduced fertility and germline defects, suggesting this gene is necessary for

proper development of the male germline. I investigated its role in germline sexual identity and found that it promotes male sexual identity in the germline. Strikingly, CG15930 is able to induce female germ cells in a male soma to adopt male fate, and go through spermatogenesis. *CG15930* makes a tudor domain protein, but its function appears to differ from that of its mouse homolog, TDRD5, which is also required male-specifically in the mouse germline.

Lastly, I also conducted a screen for genes with sex-specific functions in the germline. This led to the investigation of the role of *discs large* (*dlg*) in the male germline. Interestingly, loss of *dlg* from the germline results in severe germline depletion in males but not in females, suggesting that it has a sex-specific germline function. Further analysis revealed that *dlg* is required for germ cell proliferation and to maintain junctional integrity in the male germline.

Altogether, the work presented in this dissertation increases our understanding of germline sexual identity and sex-specific germ cell development. It provides valuable insight into how SXL acts to promote female identity in the germline, and reveals new avenues to study the mechanistic control of germline sexual identity.

Advisor: Mark Van Doren

Second Reader: Robert Johnston

**Acknowledgements**

I would like to thank all the people who contributed to my growth and success during graduate school. Firstly, I would like to thank my advisor Mark Van Doren for allowing me to work on interesting questions that challenged me. His breadth of knowledge and careful scientific style really helped me to develop my scientific acumen. I'd like to thank all of the past lab members who mentored me when I first joined the lab. I learned so much from them and I am very grateful for the time and effort they put into helping me get off to a great start. I'd especially like to thank Sheryl Southard for patiently teaching me the basics of flywork, and Shyenne Yang for her thoughtful advice. Special thanks to Kelly Baxter for being a cloning genie and an absolute joy to work with, and Erin Jimenez for being a great friend, labmate, and colleague.

Next I'd like to thank the graduate program for creating a great atmosphere in which to learn science. Thanks to Xin Chen's and Bob Johnston's labs for many many hours of feedback and lots of shared reagents. Thank you to all of my past and current thesis committee members for their advice, guidance and encouragement. Special thanks to Nick Ingolia who was instrumental during the planning of the RNA-Seq experiment, and Fred Tan for teaching 'all things Bioinformatics'. And a very special thank you to the administrative staff for doing all they can to make grad school go by smoothly.

Lastly I'd like to express heartfelt hugs and gratitude to all of my loved ones for their caring support, laughter, dancing, and delicious foods that kept me going throughout the years. I'm very lucky to have you all!

# Table of Contents

# Index of Tables and Figures

## Chapter 1: Introduction



## Chapter 2: Finding targets of Sex lethal in the germline

**Chapter 3: tudor5-prime promotes male identity in the germline**

**Chapter 5: *dlg* is required for proper development of the male germline**

# Chapter 1:

# Introduction

## Introduction

Sexual reproduction is important to maintain genetic variation within populations. Gamete production is key for sexually reproducing species. This occurs in the gonad, a highly sexually dimorphic organ. Gonads are made up of two primary types of cells: germ cells (germline) and somatic cells (soma). Both of these cell types are essential in order for proper gametogenesis to occur. Proper gametogenesis also requires successful germline stem cell maintenance, effective germline–soma communication and correct sex determination.

Sex can be determined by several different mechanisms, including sex chromosome constitution, temperature, and social cues. In *Drosophila*, sex chromosome constitution is the method used for sex determination. This decision is made in many parts of the body. But while the body is composed of one primary cell type, somatic cells, gonads also have germ cells. Both germ cells and somatic cells need to know their sexual identity in order for proper gametogenesis to occur. In some species, the sex of the soma determines the sex of the germline. In *Drosophila*, however, the germline determines its sex differently from the soma. The way in which this occurs is poorly understood. The work in this dissertation aims to improve our understanding of *Drosophila* germline sex determination. To accomplish this, I focus on refining the function of *Sex lethal*, the key gene promoting female identity in the germline.

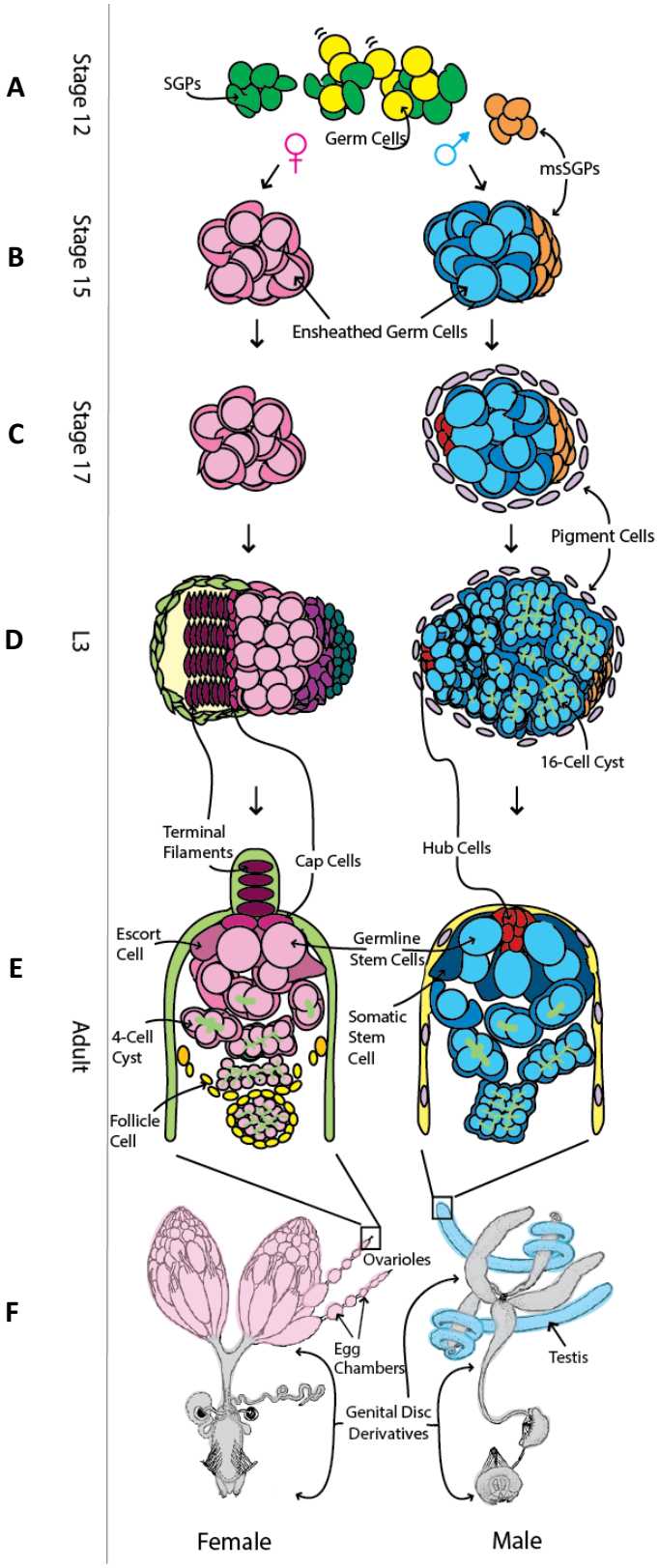
## Gonad Formation

Early in *Drosophila* embryogenesis, germ cells and somatic cells come together to form the gonad. This begins at embryonic stage 12 (Fig 1.1A), and by embryonic stage 17, germ cells and somatic cells have fully coalesced into a functional gonad (Fig .1.1C). At this stage, gonad appearance is already sex specific; testes have two populations of somatic cells that are not present in ovaries. The first population is the male-specific somatic gonadal precursors at the posterior of the testis. The second is a compact structure at the anterior of the testis known as the hub, a part of the testis niche. The niche is a special microenvironment that functions to maintain the stem cell identity of germline stem cells. In a properly formed gonad, the somatic cells send out cytoplasmic extensions to enclose the germ cells. This process is known as ensheathment (Fig 1.1B). Proper ensheathment promotes soma–germline communication, which is essential for proper gametogenesis (reviewed by Jemc, 2011).

Following gonad formation, male germ cells increase proliferation and spermatogenesis begins. From this point until adult, there is a steady increase in testis size as germ cells proliferate, enter meiosis, and differentiate into sperm (Bate & Arias, 1993). When an adult male ecloses, each testis is a long tube-like structure with an open posterior end where mature sperm is released. In contrast, ovaries remain relatively quiescent following gonad formation; ovarian

germ cells divide only three times until the third larval instar stage (L3) (Gilboa & Lehmann, 2006). Ovary morphogenesis begins in earnest at the late L3 to early pupal stage, when somatic cells intercalate to form a stack of cells called terminal filaments, an important component of the ovarian niche (Lin & Spradling, 1993). Another component of the ovarian niche are cap cells (Fig 1.1 D, E). These cells form at the base of terminal filaments, and they anchor the germline stem cells through E-cadherin mediated attachment to the niche (Song et al., 2002). In addition to niche formation, individual ovarian units called ovarioles begin to form (Fig 1.1D). When an adult female ecloses, each ovary is made up of 16-20 identical ovarioles, and each ovariole is made up of an anteriorly localized compartment called the germarium, followed by a series of egg chambers that bud off from the germarium and differentiate into eggs (Fig 1.1F).

**Fig 1.1 Schematic of Gonad Formation.** A) Gonad formation begins at embryonic stage 12 when the somatic gonadal precursors (SGPs) begin to intermingle with the germ cells. B) By stage 15, germ cells and somatic cells have coalesced. Germ cells (light blue and light pink) are ensheathed by somatic cells (dark blue and dark pink). C) By stage 17, functional gonads have formed. D) By L3, ovary morphogenesis is beginning. In the testis, germline differentiation has progressed through all the mitotic stages. E) The anterior tip of the gonads is where gametogenesis begins. Important cell types are labelled. F) Adult ovaries and testes attached to the genital tract. Ovaries are made up of identical units called ovarioles. Testes are coiled at the posterior end where sperm is released. (Illustration modified from Camara, Whitworth & VanDoren, 2008).


A
Stage 12
SGPs
Germ Cells
msSGPs
B
Stage 15
Ensheathed Germ Cells
C
Stage 17
Pigment Cells
D
L3
16-Cell Cyst
Terminal Filaments
Cap Cells
Hub Cells
E
Adult
Escort Cell
Germline Stem Cells
Somatic Stem Cell
4-Cell Cyst
Follicle Cell
F
Ovarioles
Egg Chambers
Testis
Genital Disc Derivatives
Female
Male

## Gametogenesis

In ovaries and testes, gametogenesis begins at the anterior tip of the gonads. There, distinct somatic structures called niches, maintain the germline stem cells responsible for continuously replenishing the germ cell population. Important components of the ovarian niche, the terminal filament cells, cap cells, and escort cells are present in every ovariole. In testes, the niche is made up of the hub as well as the surrounding somatic cells known as somatic stem cells (Fig 1.1E).

Oogenesis begins with the asymmetric division of a germline stem cell (GSC). This occurs in the germarium, the most anteriorly localized compartment of each ovariole (Fig 1.1E). Each germarium contains 2 to 3 GSCs anchored to the ovarian niche through E-cadherin-mediated attachment (Song et al., 2002). The asymmetric division of a germline stem cell produces one daughter cell that remains attached to the niche, retaining its stemness, and another daughter cell that moves away from the niche and begins to differentiate. This daughter cell is known as the cystoblast. The cystoblast turns on the germ cell differentiation program through the activation of the *bag of marbles* (*bam*) gene (Mckearin & Spradling, 1990). The differentiating cystoblast undergoes four rounds of mitotic divisions, each with incomplete cytokinesis. This leads to the creation of 2, 4, 8, and 16 interconnected germ cells known as an oogonial cyst. Owing to their

interconnected nature, these differentiating cysts are often simply referred to as 2, 4, 8, and 16 cell cysts (Fig 1.1 D-E).

Germ cells in a cyst are connected by intercellular bridges called ring canals. Hu-li tai shao (HTS), an adducin-like protein, is an important component of ring canals (Robinson, Cant, & Cooley, 1994). It is also an essential component of a membranous cytoplasmic structure—the fusome—which extends through ring canals, into all the germ cells in a cyst (Lin, Yue, & Spradling, 1994). The progression of cyst differentiation can be monitored by using HTS immunostain to identify the different stages of fusome structure. In cystoblasts, its structure is in the form of a circular cytoplasmic organelle known as the spectrosome. In dividing cysts, however, HTS shows the increasingly elongating fusome structure, which branches into each germ cell of the cyst.

Following the mitotic divisions, oocyte specification takes place. In ovarioles, a single germ cell from each 16 cell cyst is chosen to become the oocyte. Typically, the germ cell with four ring canals and the most fusome material becomes the oocyte (de Cuevas & Spradling, 1998; reviewed by Huynh & St Johnston, 2004). The remaining 15 germ cells become nurse cells, which nurture the oocyte throughout its development. Following oocyte fate specification, the oogonial cyst becomes encapsulated by a layer of somatic cells known as follicle cells (Fig 1.1E yellow cells), and buds off from the germarium as an egg chamber.

Successive stages of egg chambers gradually increase in size as the oocyte develops. During this time, the oocyte enters meiosis and eventually arrests at metaphase I. In contrast, the nurse cells undergo numerous cycles of DNA synthesis without mitosis, thereby becoming polyploid. In this way, they can produce an abundance of mRNAs and proteins that are transferred to the growing oocyte. These maternally-deposited nutrients are important to fuel the rapid cell cycles at the beginning of embryogenesis.

Spermatogenesis is very similar to oogenesis. It begins with the asymmetric division of a germline stem cell (GSC), attached to the hub/niche at the anterior tip of the testis. Each testis has more GSCs than an ovariole, however, 10 to 12 GSCs surround a hub. Following asymmetric GSC division, the gonialblast daughter cell leaves the niche environment and initiates differentiation. Similar to ovaries, each gonialblast undergoes 4 rounds of incomplete mitoses forming 16-cell interconnected spermatogonial cysts. Unlike ovaries however, every germ cell in a spermatogonial cyst enters meiosis to become a spermatocyte and produce sperm. Meiotic entry begins with a period of growth and gene expression, during which each spermatocyte expresses testis-specific TBP-associated factors (tTAFs), essential for the activation of spermatid differentiation, and completion of the meiotic cell cycle (Lin et al., 1996; White-Cooper et al., 1998; Hiller et al., 2004). Each spermatocyte produces 64 round

spermatids, which elongate and differentiate into mature sperm (reviewed in Fuller, 1998). Therefore, in males, each germline stem cell division produces 64 mature sperm.

## Germline Stem Cell Maintenance

Germline stem cells are the progenitor cells of all differentiating germ cells. Therefore, it is crucial that germline stem cells are properly maintained in their niche. In males, the JAK/STAT pathway plays a major role in GSC maintenance. JAK/STAT signaling from the niche to the GSCs upregulates STAT92E in GSCs. Proper STAT92E upregulation is important to maintain E-cadherin-mediated GSC attachment to the hub (Leatherman and Dinardo, 2010).

In females, the main signaling pathway used for GSC maintenance is the Bone morphogenetic protein (BMP) pathway. The BMP ligands Decapentaplegic (DPP) and Glass bottom boat (GBB) produced from the niche (Xie & Spradling, 1998) activate the BMP signaling pathway in germline stem cells. BMP signaling activates the *Drosophila* Smads, *Mad* and *Medea*, in GSCs. These genes achieve stem cell maintenance by repressing transcription of the differentiation gene *bag of marbles* (*bam*) (Song et al, 2004).

## Somatic Sex Determination

Sex determination is an essential process for sexually reproducing organisms. The germ cells and somatic cells in *Drosophila* gonads, both need to "know" their sex, but they make this decision in different ways. Sex in both germline and somatic cells is determined by the X chromosome dose (Schupbach, 1985, Steinmann-Zwicky et al, 1989; Erickson & Quintero, 2007). In the soma, an XX dosage signal targets *Sex Lethal*'s (*Sxl)* early promoter, *Sxl*-Pe, which turns on transiently only in female embryos (Fig 1.2A). SXL activates an alternative splicing cascade culminating in the expression of female-specific isoforms of the somatic sex regulators, *doublesex* (*dsx*) and *fruitless* (*fru*). In males (XY), a single X chromosome dose is not sufficient for *Sxl*-Pe activation, and as a result, *doublesex* and *fruitless* transcripts are spliced in the default splicing pattern, leading to the expression of male specific proteins (reviewed by Camara et al, 2008).

The RNA binding protein SXL is the key regulator of sex in the soma. To accomplish this important task, *Sxl*'s regulation is tightly controlled at different levels. The first layer of control is at the transcriptional level, with the activation of the *Sxl* early promoter only in females. And the second layer of control is at the splicing level. The SXL protein made from the early promoter is required to mediate alternative splicing of its own pre-mRNA made from the late, non sex-specific promoter (*Sxl*-Pm) (Keyes et al, 1992). In this way, *Sxl* transcript

produced from the late promoter undergoes SXL-mediated splicing, to produce a functional SXL protein in females. Default splicing of the *Sxl* transcript occurs in males. The default-spliced transcript contains an exon with a premature stop codon. Therefore, this message translates into a truncated, nonfunctional SXL protein (Fig 1.2B).

The second gene in the somatic alternative splicing cascade is *transformer* (*tra*). SXL-mediated *tra* splicing occurs in a manner very similar to the splicing of *Sxl* itself. SXL protein blocks a splice site, which results in the splicing out of an exon containing a premature stop codon. In this manner, functional TRA protein is produced in females but not in males (Fig 1.2C). Transformer is the direct regulator of $dsx^F$ production in females. This also occur through an alternative splicing mechanism (Figure 1.2). However, the splicing mechanism used by TRA is different from that of SXL. TRA, in collaboration with its partner TRA2, enhances the use of a splice site resulting in a female-specific spliceform of *dsx*, which creates a female-specific DSX protein (Figure 1.2 D-E) (reviewed by Camara et al, 2008).

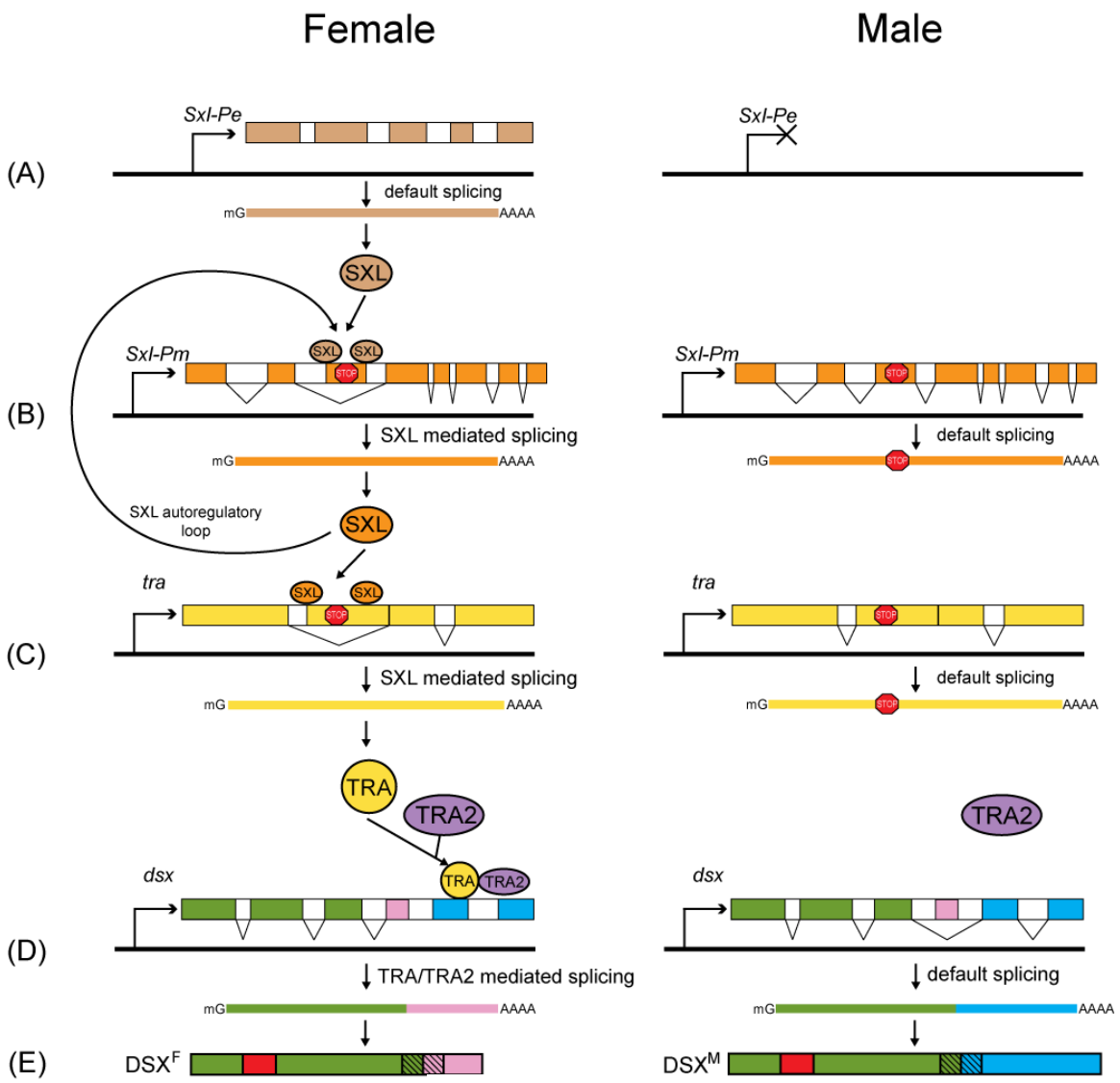



**Fig 1.2: Schematic of the somatic sex determination alternative splicing cascade.** A) *Sxl* expression from the early promoter occurs only in females. B) The *Sxl* auto-regulatory loop. SXL protein from the early promoter splices *Sxl* transcript from the late promoter (*Sxl*-Pm). Default *Sxl* splicing occurs in males C) SXL directs alternative splicing of its downstream target *tra*. D) TRA and TRA2 splice *dsx* into E) female and male-specific transcripts. (Illustration modified from Camara, Whitworth & VanDoren, 2008).

## Germline Sex Determination

The germline sex determination pathway is not as well understood as its somatic counterpart. The soma has been shown to contribute to germline sexual identity. In males, JAK/STAT signaling from the soma to the germline is important to maintain the sexual identity of male germ cells (Wawersik et al, 2005). And in both sexes, the sex of the soma affects the transcription of certain genes in the germline (Hinson & Nagoshi, 1999; Casper & Van Doren, 2009). But somatic control of germline sexual identity is limited, as germ cell sexual identity is not reversed while in the soma of the opposite sex. Male germ cells transplanted into an ovary are unable to go through oogenesis, neither can female germ cells in a testis undergo spermatogenesis (schupbach, 1982; Schupbach, 1985; Steinmann-Zwicky et al, 1989). This indicates that germ cells have some autonomous control over their own sexual identity—X chromosome dose is also important for germ cell sex determination.

*Sex lethal* is required in the female germline both for proper oogenesis and for germline sexual identity (Schupbach, 1985; Steinmann-Zwicky et al, 1989; Bopp et al, 1999). However, experiments using germ cell mutants of the downstream somatic targets of SXL, showed that the germline sex determination pathway differs from that of the soma, as neither *tra* nor *dsx* is required in the germline (Marsh & Wieschaus, 1978; Schupbach, 1982). *Sxl's* requirement in the

germline is cell autonomous; mutants of *Sxl* that affect its function only in the germline impair female germline development. These mutants produce a wild type somatic ovary filled with undifferentiated germ cells, the "ovarian tumor" phenotype (Salz et al, 1987). The genes, *ovarian tumor* (*otu*) and *ovo*, have also been identified as involved in female germline sex determination (Oliver et al, 1990; Oliver et al, 1993), as mutants of these genes phenocopy the *Sxl* germline mutant phenotype. These genes are thought to act upstream of *Sxl* (Oliver et al, 1990; Oliver et al, 1993), but much about their mechanism of action remains unknown. It has been shown, however, that *Sxl* plays a dominant role promoting female identity in the germline (Hashiyama et al, 2011). This group showed that while a wild type male germ cell cannot go through oogenesis if transplanted into an ovary, when a male germ cell ectopically expressing *Sxl* is transplanted into a female ovary, it completes oogenesis and produces fertile offspring. This was the strongest evidence to date showing the significance of SXL function for female identity in the germline. But we still did not understand how SXL was doing this, or know what its germline targets were.

Before the work presented in this dissertation was done, the only known germline target of SXL was *nanos*. But *nanos* is not involved in sexual identity; instead, it is important for germline stem cell maintenance (Chau et al, 2012). This group showed that SXL in cooperation with BAM translationally represses

*nanos* to allow cystoblast differentiation. The work presented in this dissertation proposes many possible germline targets of SXL, and discusses two verified targets of SXL; one was previously characterized by another member of our lab (Yang et al, 2012), and the other is characterized within.

# Chapter 2:

# Finding targets of Sex lethal in the germline

## Introduction:

*Drosophila* germline sex determination is not well understood. We do know a few things about it however. We know that signaling from the somatic gonad to the germline contributes to germ cell sexual identity (Hinson and Nagoshi, 1999; Wawersik et al, 2005). This was well shown by work from a previous student in our lab. She found that masculinization of an XX soma is sufficient to induce XX germ cells to express male-specific genes, and a feminized soma can repress expression of male genes in XY germ cells (Casper & Van Doren, 2009). However, germline transformation is incomplete in the somatic environment of the opposite sex, as these germ cells cannot form viable gametes. This suggests that germ cells are inherently sexually dimorphic, and have a germ cell autonomous contribution to determining their own sexual identity.

The RNA-binding protein *Sex lethal* (*Sxl*) is the master sex switch for somatic sex determination. SXL protein is functional only in females, and in the soma it acts as a regulator of both alternative splicing and translation. Specifically, it can bind to uridine-rich regions in introns to mediate alternative splicing of its targets. It can also bind to the 3'UTR to repress translation of a target. In the soma, *Sxl* controls sex determination by initiating an alternative splicing cascade which ultimately leads to the expression of female or male

isoforms of *doublesex* and *fruitless*, the key regulators of sex-specific gene expression in the soma.

*Sxl* has long been known to also be important in the female germline for germline sexual identity (Schupbach, 1985, Staab et al, 1996). How SXL regulates germline sexual identity, however, is still unknown other than the fact that SXL's downstream targets in the soma, *transformer* and *doublesex*, are not important in the germline. A pivotal study published in 2011 refined *Sxl*'s germline role (Hashiyama et al, 2011). This study showed that while an XY germ cell transplanted into an ovary cannot differentiate and will instead form a tumor, a transplanted XY germ cell ectopically expressing *Sxl* can go through normal oogenesis and even produce fertile offspring (Fig 2.1). This is strong evidence showing that *Sxl* acts autonomously in female germ cells as the master gene directing female identity in the germline.

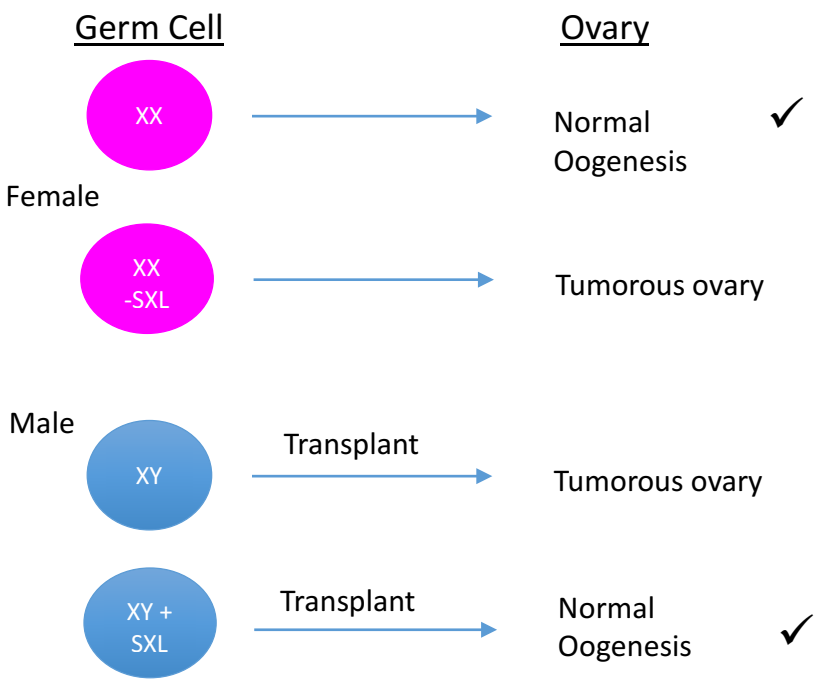



**Fig 2.1: *Sxl* plays a dominant role promoting female germline sex identity**

Female germ cells can only go through oogenesis if *Sxl* is expressed in the germ cells. Female germ cells that are mutant for *Sxl* (XX -*Sxl*) are unable to complete oogenesis; instead, these germ cells cause an ovarian tumor. Male germ cells, transplanted into an ovary also cause a tumor. However, the ectopic expression of *Sxl* in male germ cells (XY + *Sxl*) gives them the ability to successfully complete oogenesis. *Sxl expression is sufficient to induce female development in XY germ cells.*

To understand how *Sxl* directs female identity in the germline, it is essential to discover its target genes in the germline. In this chapter, I will discuss an RNA expression profiling (RNA-Seq) experiment that I conducted and analyzed in an effort to find targets of *Sxl* in the germline. The RNA-Seq was conducted using ovaries that lack SXL function in the germline, compared to ovaries with SXL function. Analysis of this data showed that only 200 genes were differentially expressed between these two samples. 23 percent of these genes showed very large expression differences of 8 fold and higher. I selected one of these highly expressed genes, CG15930, to conduct an in-depth study of its function and discovered that it is indeed a target of *Sxl* and involved in germline sexual identity. Specifically this gene promotes male identity in the germline and is repressed by SXL in the female germline (discussed fully in Chapter 3). SXL-mediated repression of target genes appeared to be a theme in the RNA-Seq, as more differentially expressed genes were upregulated in the *Sxl*-RNAi sample compared to the control. This suggests that these genes are normally repressed by SXL function in the ovary. In addition to whole gene changes, differential exon usage was also analyzed. This analysis produced many interesting genes with alternatively spliced exons that may be direct or indirect targets of SXL. This RNA-Seq data and the studies that will spring from it will make a

significant contribution to our understanding of how *Sxl* functions to promote female sexual identity in the germline.

## Materials and Methods

### Fly stocks

The fly stocks used were obtained from Bloomington Stock Center unless otherwise indicated. *bam*$^{1}$ (McKearin & Spradling, 1990), *bam*$^{\Delta 86}$ (BDSC# 5427), *nos*-Gal4 (Van Doren et al, 1998), the control RNAi used was p{VALIUM20-mCherry}attP2 (BDSC# 35785), uas-*Sxl*-RNAi=TRiP.HMS00609 (BDSC# 34393), *bam*:HA was obtained from Xin Chen (Eun et al, 2013).

### Immunofluorescence

Adult ovaries and testes were fixed, blocked and stained as described in Gonczy et al, 1997. All images were taken with a Zeiss LSM 510 confocal microscope. Primary antibodies and the concentrations used are as follows: chicken anti-Vasa 1:10,000 (K. Howard); rabbit anti-Vasa 1:10,000 (R. Lehmann); mouse anti-SXL 1:8 (M18, DSHB); mouse anti-armadillo 1:100 (N2 7A1, DSHB); rat anti-HA 1:100 (3F10, Roche); guinea pig $\alpha$-TJ (1:1,000; generated by J. Jemc using the same epitope described in Li et al, 2003). DSHB: Developmental Studies Hybridoma

Bank. Secondary antibodies were used at 1:500 (Alexa-fluor). Stains were mounted in vectashield mounting solution with DAPI (vector Industries).

## Library Generation and sequencing

Gonads were dissected from 1-3 day old flies raised at $25^0$C. Ovaries were dissected from virgin females. 3 biological replicates were dissected for each genotype. Total RNA was isolated from all genotypes using RNA-bee (Tel-Test). Contaminating DNA was removed from the RNA using Turbo-DNA-free (Ambion). 200ng of RNA was used to prepare each library using the illumina TruSeq RNA Library Prep kit v2. 100bp paired-end read sequencing was done by the Johns Hopkins Genetic Resources Core Facility. The *bam* mutant male and female libraries were sequenced in one lane and the *Sxl*-RNAi, control-RNAi libraries were sequenced in separate lane, therefore having 6 libraries per lane.

## Read mapping and quality

Quality of raw reads was assessed using the fastQC kit (Babraham Institute). RNA-Seq reads were mapped to the *Drosophila* genome using TopHat2 version v2.0.9 (Trapnell et al, 2012, Kim et al, 2013) and *Drosophila* Ensembl annotation BDGP5 with options –i 42 and –g 1.

Differential gene expression and differential exon expression analysis

Differential gene expression analysis was done using Cufflinks cuffdiff v2.1.1 (Trapnell et al, 2010), using ensemble annotation BDGP5 with options –b and --max bundle frags 500,000. Adjusted P value of <0.05 used for significance cutoff. Cuffdiff data was analyzed using Cummerbund (Trapnell et al, 2012). Annotation building was done using Cufflinks cuffmerge (Trapnell et al, 2010), using the Galaxy tool version 0.0.6. Differential exon analysis was done using DEXSeq (Anders et al, 2012).

## Results

Setup of the RNA-Seq experiment

RNA-Seq libraries were generated in biological triplicate for each genotype, and 100bp paired-end reads were generated from each replicate. These were the genotypes used to generate data for the *Sxl*-RNAi RNA-Seq experiment: both the control and the experimental genotypes were in the *bag of marbles* (*bam)* mutant background; the control genotype was expressing a control-RNAi in the germline, and the experimental genotype was expressing *Sxl*-RNAi in the germline. Use of the *bam* mutant as the background for this RNA-Seq experiment gave 3 distinct advantages. Firstly, *bam* is a gene that is required for germline

differentiation; it is expressed in the differentiating cystoblast in the ovary (Fig 2.2A), and turns on the germline differentiation program (McKearin and Spradling, 1990). In *bam* mutants germ cells cannot differentiate, so the early germ cells continue to divide and fill up the ovaries with undifferentiated germ cells. Because of this, the expression profiles obtained from these tissues are indicative of genes present in the early (undifferentiated) germline. And the undifferentiated germline are the population of germ cells we were interested in studying. This is where SXL protein is expressed in wild type ovaries (Fig 2.2B, arrows). This leads to the second advantage to using *bam* mutants for this RNA-Seq experiment. Since the *bam* mutant ovary is completely populated with undifferentiated germ cells, all of these germ cells are expressing SXL (Fig 2.2C). This provides a highly robust comparison of SXL function; comparing ovaries with expression of *Sxl* all throughout the germline to ovaries with absolutely no expression of *Sxl* in the germline (compare Fig 2.2 C and D).

The third advantage of using *bam* mutants is that it gives us the ability to compare similar types of tissues. Since *Sxl*-RNAi also causes a tumorous phenotype, doing the control-RNAi in the *bam* mutant background gives the ability to compare a *bam* control-RNAi tumor to a *bam Sxl*-RNAi tumor. The alternative—comparing *Sxl*-RNAi tumorous ovaries to wild type ovaries— is not a fair comparison. This is because wild type ovaries have many stages of

differentiating egg chambers that are absent from tumorous ovaries (Fig 2.2B', arrowheads and Fig 2.3). Differentiating egg chambers turn on a unique gene expression profile indicative of differentiating germ cells, and nurse cells have dramatically increased levels of transcription (explained in chapter 1, page 9). In this scenario, the differential gene expression observed between the 2 genotypes would be artificially high reflecting the difference in differentiation stages present, instead of differences caused by loss of *Sxl*.

A companion RNA-Seq dataset was generated using male and female adult gonads from *bam* mutant flies. This dataset serves two purposes: first, the *bam* mutant female dataset serves as a second control for the *Sxl*-RNAi data. Secondly, the *bam* mutant male dataset serves as way to find genes in the *Sxl*-RNAi dataset that are not just differentially expressed, but whose expression pattern have changed to reflect the male expression pattern. There are several genes from the dataset that fit this criteria.

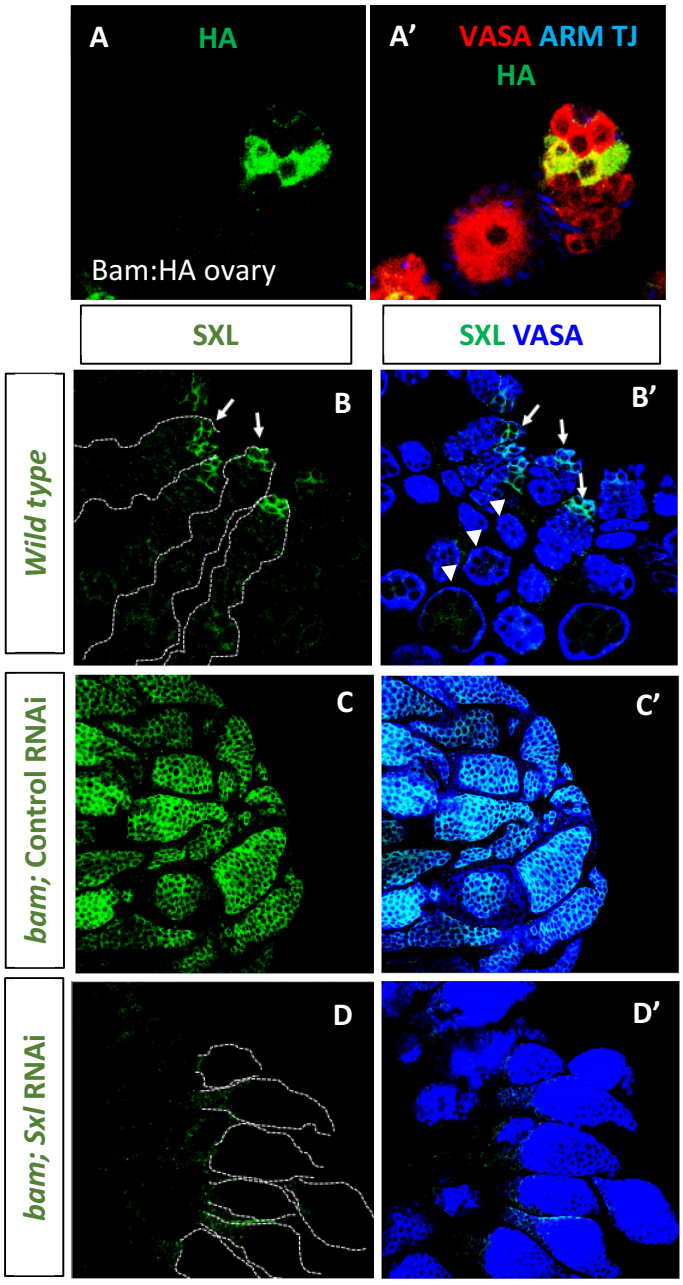



**Fig 2.2: Expression of SXL in ovaries**. A) Confocal image of an adult ovariole expressing a bam:HA fusion construct. B-D) Confocal images of adult B) wild type ovaries, C) *bam* mutant control-RNAi ovaries, and D) *bam* mutant *Sxl*-RNAi ovaries showing expression of SXL in all. Dashed lines outline ovarioles. Arrows mark the tip of ovarioles. Arrowheads mark differentiating egg chambers.

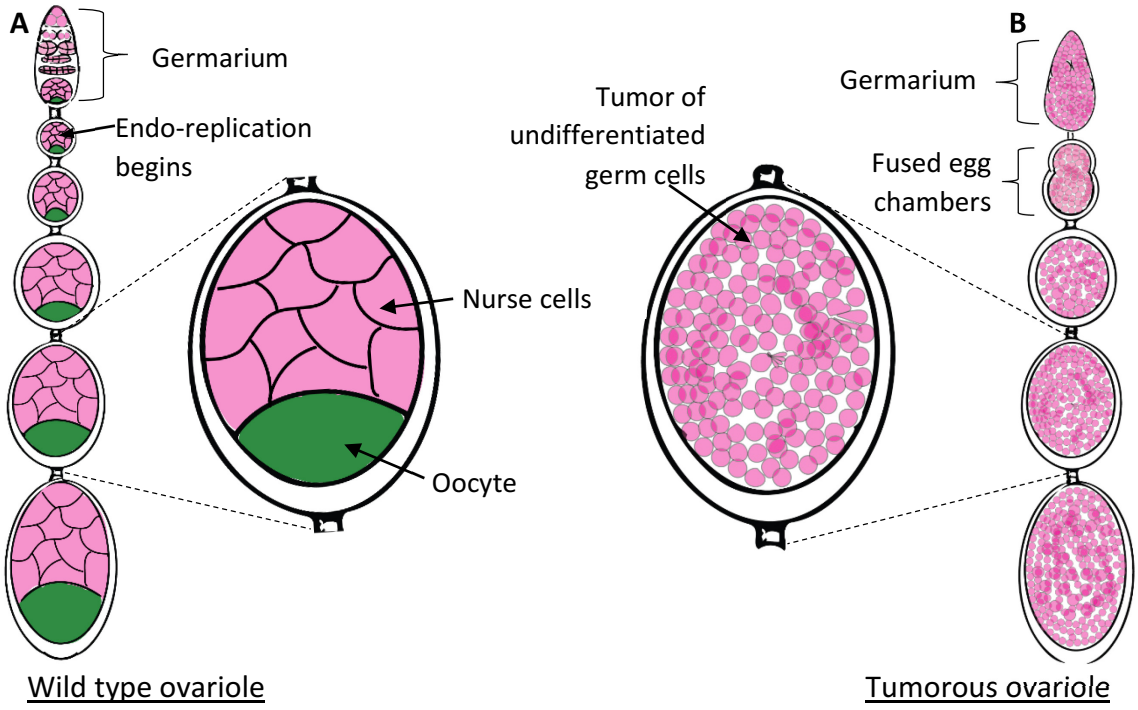



**Fig 2.3: *Sxl* mutant ovaries are tumorous**. A) Illustration of a wild type ovariole. Germ cells are shown in pink. Early/undifferentiated germ cells, such as germline stem cells, reside in the germarium. Egg chambers increase in size as germ cell differentiation proceeds. Nurse cells endoreplicate and increase transcription and translation to nurture the developing oocyte (green).
B) Illustration of a tumorous ovariole, caused by the over-proliferation of the early/undifferentiated germ cells normally restricted to the germarium. Tumorous ovaries can result from mutations in different genes including *bam* and *Sxl*. A common developmental defect in tumorous ovaries are fused egg chambers. Enlarged portions show the dramatic difference between the egg chambers of wild type ovaries and tumorous ovaries.

## The RNA sequencing data generated was of high quality

The quality of RNA-Seq data significantly affects the accuracy of the data analysis. To determine whether the raw RNA-Seq reads were of high quality, I used the fastQC RNA quality check kit. This program and others like it uses different metrics all designed to assess quality of the raw reads and detect possible contamination in a sequencing library. FASTQC determined that the raw reads for all the libraries were of very high quality—over 95% of the raw reads received a high quality score. Because of this, there was no need to pre-process the reads (by trimming to remove low quality bases for example), and these raw reads were used directly for mapping. Other than having a high level of duplication, the raw reads passed all quality check categories used to ensure good quality of RNA-Seq data (Fig 2.4).

As expected for high quality reads, the percentage of reads that mapped to the *Drosophila* genome was very high. Each library had more than 85% of reads uniquely mapping back to the *Drosophila* genome (Table 2.1). Therefore this data gave very high coverage of the *Drosophila* genome. The *Sxl*-RNAi experiment alone produced ~150 million reads. The *Drosophila* genome is about 30Mbps, which means that this experiment provides an astounding 500X coverage of the *Drosophila* genome.

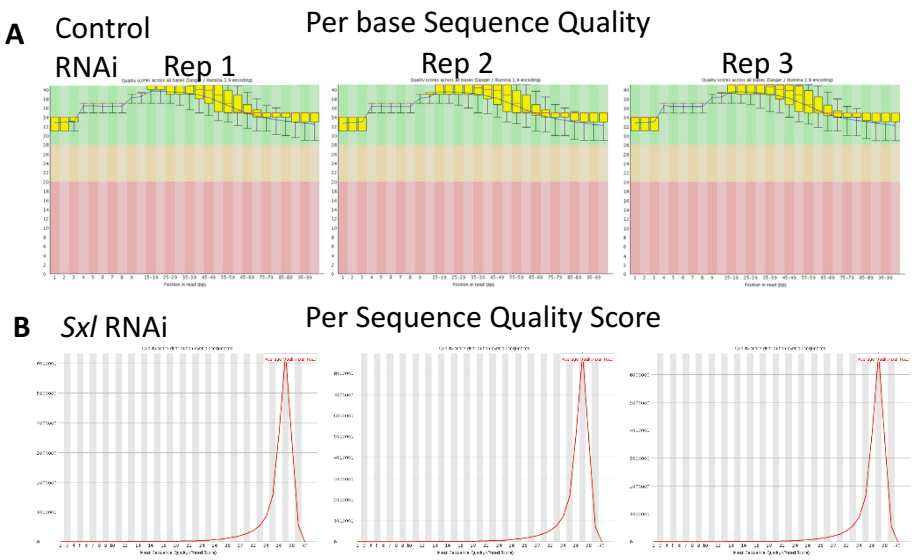



**Fig 2.4: High quality of the raw RNA-Seq reads**. A) Representative per base sequence quality graphs. Graph represents an average sampling of the raw reads in each fasta file. Y-axis shows the quality scores. The background is colored so that green represents very good quality, orange represents reasonable quality and red represents poor quality. X-axis shows base position. A box-whisker plot is drawn at each base position to represent the variation of quality of the bases at that position in all reads samples. B) Representative per sequence quality score graphs for a sampling of the raw reads in specified fasta file. Y-axis shows number of reads. X-axis shows the quality scores. The highest possible Q-score is 38. Q-scores of 30 or more are considered high quality. These graphs are representative for all replicates of the *Sxl*-RNAi RNA-Seq experiment as well as the companion *bam* RNA-Seq experiment.

| sample | Total # of reads | % uniquely mapped |
|---|---|---|
| *Sxl*-rep1 | 18,606,954 | 86.6<br>85.9 |
| *Sxl*-rep2 | 26,164,948 | 87.4<br>86.7 |
| *Sxl*-rep3 | 19,575,554 | 88.4<br>87.6 |
| Control rep1 | 28,444,487 | 86.7<br>85.9 |
| Control rep2 | 27,122,860 | 87.2<br>86.4 |
| Control rep3 | 23,936,361 | 88.3<br>87.6 |
| *bam* male rep1 | 19,406,865 | 87.3<br>86.8 |
| *bam* male rep2 | 18,265,856 | 86.9<br>85.4 |
| *bam* male rep3 | 20,190,019 | 87.8<br>87.1 |
| *bam* female rep1 | 22,735,241 | 87.9<br>87.2 |
| *bam* female rep2 | 18,466,702 | 88.2<br>87.6 |
| *bam* female rep3 | 21,432,357 | 87.1<br>86.5 |

**Table 2.1: Mapping statistics for all RNA-Seq libraries.** Reads mapped using Tophat and Bowtie 2 to Ensembl genome annotation BDGP5**.** Percent uniquely mapped for each replicate: numbers represent unique mapping for left read and right read of the pair.

Another important quality check to validate the quality of RNA-Seq data is to ensure that certain genes used as positive controls are expressed as expected. For this quality check, I evaluated the expression of *actin5C* which should be expressed at the same relative level in all the samples. I also looked at *Sxl* which should have a dynamic expression pattern among the samples. *Sxl* expression should be high in the *bam* female control as well as the Control-RNAi, but it should be lower in the *Sxl*-RNAi (as it is knocked down in the germline though still present in the soma), and it should be even lower in males. This is exactly what the data shows; actin is expressed at similar levels in all samples, whereas *Sxl* is 3 fold lower in the *Sxl*-RNAi dataset compared to control-RNAi (Fig 2.5).

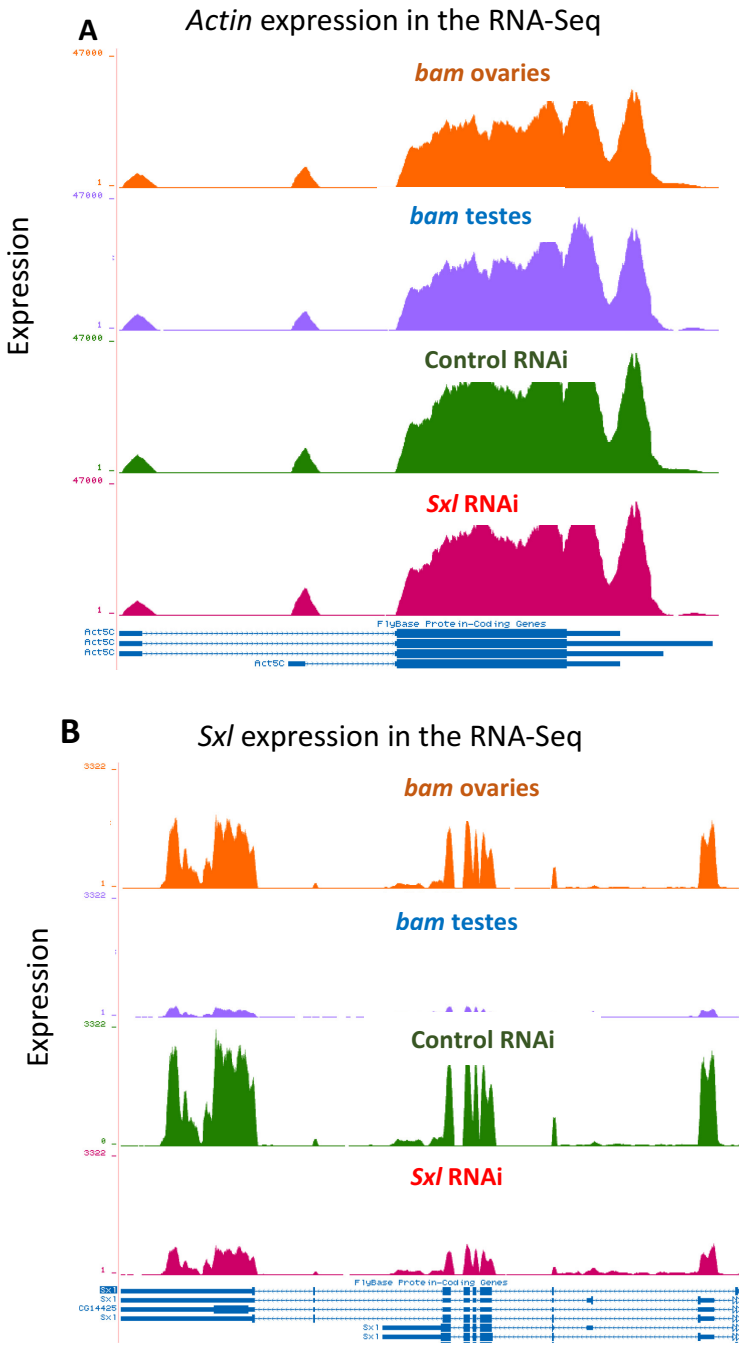



**Fig 2.5: UCSC genome browser views of positive control genes for the RNA-Seq data**. A) *Act5C* expression is similar among all datasets. B) *Sxl* expression is similar in the *bam* and control-RNAi ovaries, reduced in the *Sxl*-RNAi ovaries and very low in males.

## Differential gene expression analysis produced many interesting gene candidates

Differential gene expression analysis revealed only ~200 genes showing significant differential expression between the Control-RNAi and *Sxl*-RNAi ovaries. In contrast, thousands of genes were significantly differentially expressed in the *bam* mutant male versus female sequencing dataset. The small number of genes that are significantly different in the *Sxl*-RNAi comparison is an important indication of how similar these two tissues are. Any bioinformatics analysis is expected to produce false positives; some genes such as heat shock genes show expression that is likely artificially different between the two samples. If we remove these genes, then what remains represents a small number of hits that are indirect targets of *Sxl* at the transcriptional level, but possibly also direct targets of *Sxl* at the translational level. We are currently investigating whether this is the case for *CG15930,* the gene discussed in chapter 3.

Of the genes that were differentially expressed (Appendix A and B), a slightly larger number of genes (65%) were upregulated in the *Sxl*-RNAi sample compared to the control (Fig 2.6A). And half of this 65% are genes that have not yet been characterized, represented as CGs (Appendix A). This suggests the possibility that many of the genes that regulate germline sex are unknown because they simply have not yet been characterized.

The trend that a higher number of genes are upregulated in *Sxl*-RNAi ovaries was observed in several other instances during my analysis of the RNA-Seq data. 23% of the differentially expressed genes showed a very high differential expression—of 8 fold and higher—between the two samples (Fig 2.6E). And of this subset of genes a substantially higher percentage of them, 85%, are upregulated in the *Sxl*-RNAi sample compared to the control-RNAi, and 65% are uncharacterized genes.

The third observation of the trend is that of the genes that are differentially expressed on the X chromosome, the majority of them, again 85%, are upregulated in the *Sxl*-RNAi dataset (Fig 2.6C). This distribution of differentially expressed genes is unique to the *Sxl*-RNAi dataset. The *bam* mutant male vs female companion dataset shows a more uniform distribution of genes that are upregulated versus downregulated on each of the major chromosomes (Fig 2.6B). The significance of this observation increased with the finding that almost half of the genes upregulated on the X chromosome in *Sxl*-RNAi ovaries, are also upregulated in *bam* mutant testes (Fig 2.6D). This means that in the absence of *Sxl*, these genes adopt a male expression pattern. Interestingly, most of the genes currently known to play a role in germline sexual identity are also all on the X chromosome. These genes are *snf*, *Sxl*, *ovo* and *otu*, which all promote female identity in the germline, as well as *Phf7* which promotes male germline

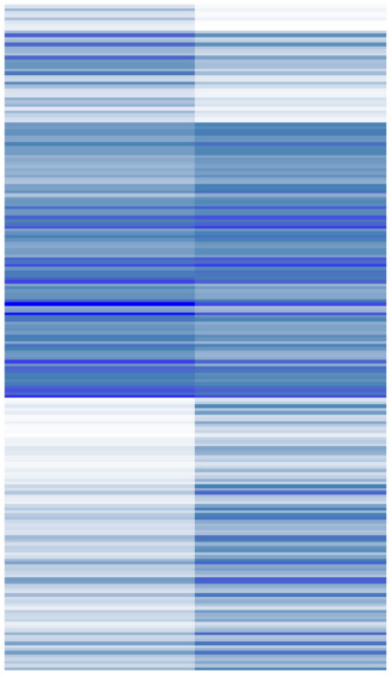
identity. Still other genes in this category are *hop*, *dome*, and *upd*, members of the JAK/STAT signaling pathway; male germ cells respond to JAK/STAT signaling from the soma to promote their sexual identity (Wawersik et al, 2005).

One surprising observation from the RNA-Seq analysis is that a number of nutrition-related genes were very highly differentially expressed, and again most were upregulated in the *Sxl*-RNAi dataset. The top 3 most highly differentially expressed genes are upregulated in the *Sxl*-RNAi sample, and 2 of these are glycolysis-related genes with 30 and 45 fold upregulation in *Sxl*-RNAi ovaries. The 2$^{nd}$ most highly differentially expressed gene is a previously uncharacterized gene, *CG42704*, which is 40 fold upregulated in *Sxl*-RNAi ovaries (Fig 2.6E). It would be interesting to determine whether *Sxl* regulates nutrition in the germline. Other interesting genes that were differentially expressed includes genes with known functions in sexual reproduction and oogenesis, germ cell migration, and splicing.

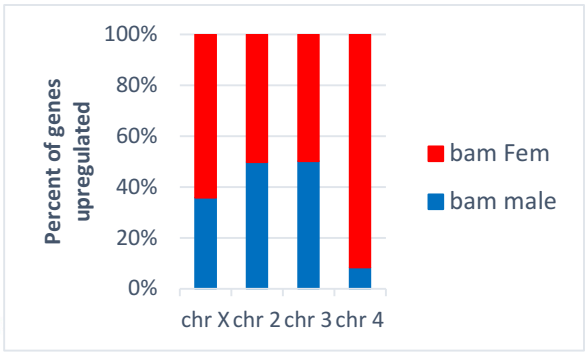
**Fig 2.6: Summary of differential gene expression analysis in *Sxl*-RNAi.** A) heatmap showing the expression of the genes that were called as significantly differentially expressed in the *Sxl*-RNAi sequencing data. B-C) The chromosome distribution of differentially expressed genes, 2 fold and higher differential expression in B) the companion *bam* RNA-Seq and the C) the *Sxl*-RNAi RNA-Seq experiment. Overlap between the genes upregulated on the X chromosome in *Sxl*-RNAi and also upregulated on the X chromosome in *bam* mutant males. The list of shared genes are shown. *Phf7*-RC is starred because only this male-specific isoform is called as significantly upregulated in *Sxl*-RNAi, not the whole gene. E) bar graph showing the expression level of genes that were 8-fold or more differentially expressed.

**A**

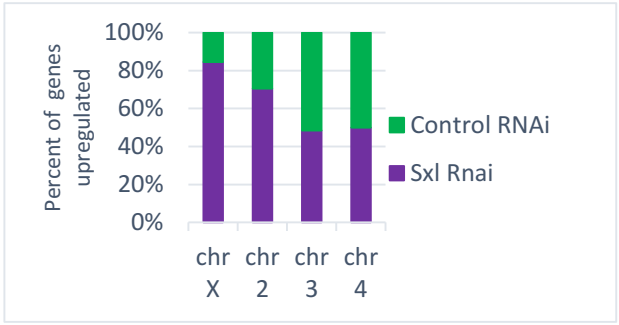
**Heatmap of all diff expr genes**



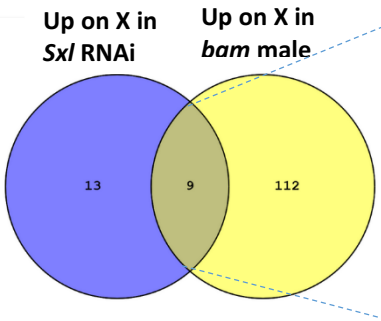

**B**




**C**

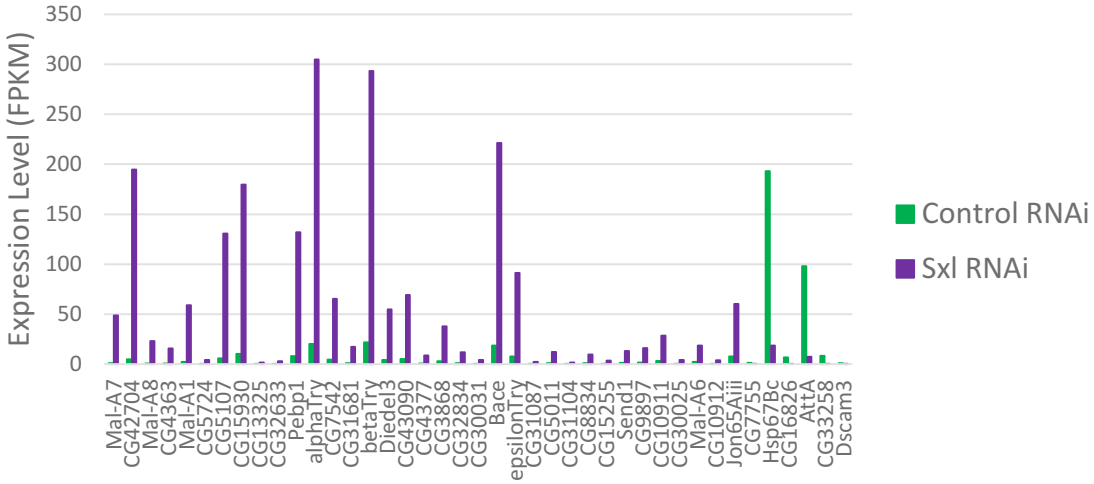



**D**




| |
|---|
| *CG15930* |
| *CG34434* |
| *dhd* |
| *CG6324* |
| *CG15599* |
| *CG32706* |
| *CG14193* |
| *CG12728* |
| *CG2865* |
| *Phf7-RC** |

**E**



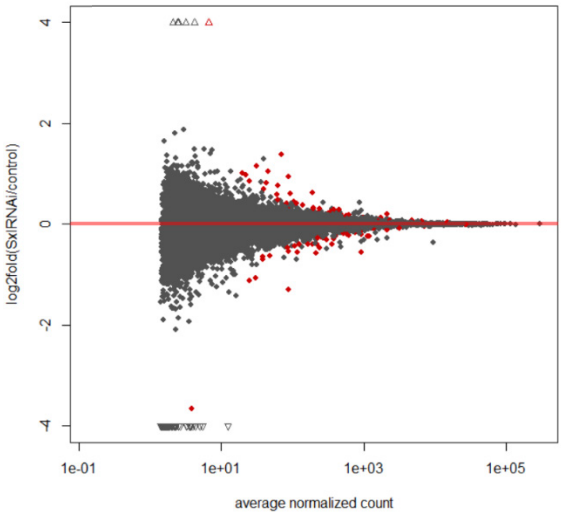
Since SXL is known to use alternative splicing to alter the transcript structure of its target genes, differential exon expression was an important analysis to conduct using this RNA-Seq data. The RNA-Seq libraries were sequenced with this is mind; using long (100bp) paired end reads to give the most information about transcript structure. This analysis was done in 2 different ways: using the publicly available *Drosophila* annotation, and using the annotation produced by my combined RNA-Seq datasets. Use of the publicly available annotation produced good results (Fig 2.7A). A total of 74 exons were determined to be differentially expressed between the 2 samples, and the expected positive controls; the *Sxl* and *Phf7* male exons were also present in the candidates list produced (Fig 2.7 C and D). However, visual analysis of the expression of the other exon candidates was not very convincing.

To maximize the use of the RNA-Seq data and improve the program's ability to find differentially expressed exons that are unannotated, I repeated the differential exon analysis using the Cufflinks-built annotation. The annotation built by Cufflinks combines the expression data from the *Sxl*-RNAi and control-RNAi datasets. This second analysis produced over 200 differentially expressed exons. The dramatic increase in identified exons highlights a wealth of transcript information added by this deep sequencing data. The new analysis discovered candidates which show differences in transcript structure that are not

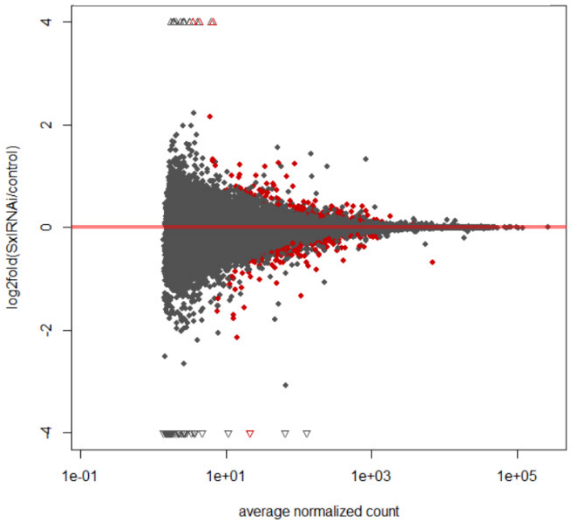
reported in the publicly available annotation. One such candidate is the gene *Clic*. *Clic* shows differential expression in the distal portion of the first exon (Fig 2.7E). This is likely due to an unannotated internal splice site in that exon. This expression pattern could also be due to the use of an alternative promoter. Use of the shorter version of this exon vs the longer version, changes the coding region of this gene, which could be biologically significant. Another feature that makes *Clic* an exciting target is that it has *Sxl* binding sites: one site 73bp into the first exon, a second site within the first intron, and another further down in the gene. The placement of the first two *Sxl* binding sites makes them optimal for SXL-mediated splicing of the first exon. Further studies into *Sxl*'s regulation of *Clic* are needed to uncover whether or not it is a direct target of *Sxl*

**Fig 2.7: Known and newly discovered differential exon usages.** A-B) MA plot of differential exon expression analysis showing mean expression on the X-axis and log2 fold change of *Sxl*-RNAi/control-RNAi on the Y-axis. Significant hits at an adjusted P-value <0.1 are colored in red for A) analysis with BDGP5 annotation and B) analysis with Cufflinks-built annotation. C-E) UCSC genome browser views of differential splicing in C) *Phf7*: expression scales are the same in each sample, D) *Sxl*: expression scales are relative to the highest expressing sample so that *bam* male expression will be visible, and E) *Clic*: expression scales are the same in both samples

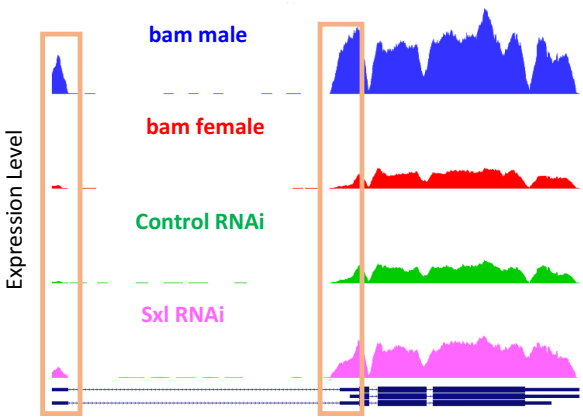
**A** Differential exon analysis using BDGP5 annotation



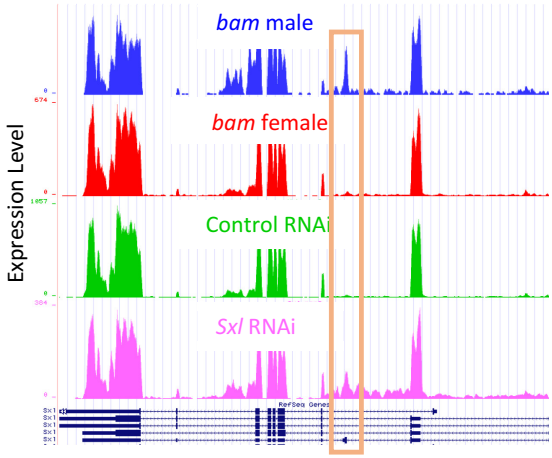

**B** Differential exon analysis using Cufflinks-built annotation



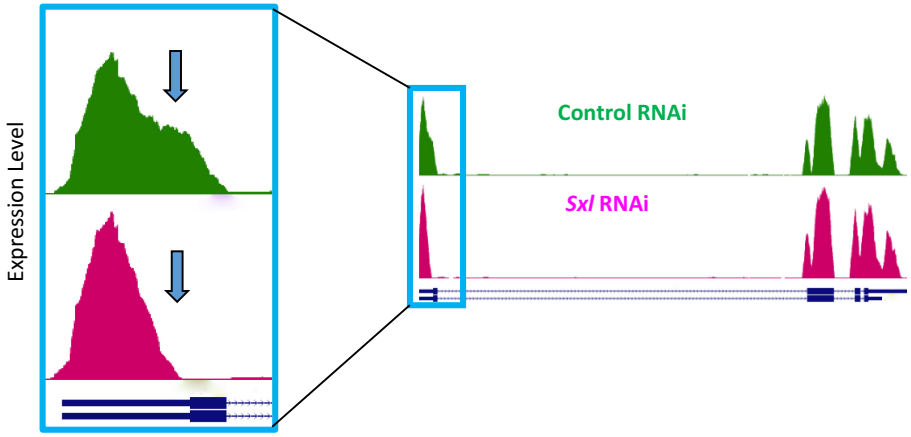

**C** Differential expression of *Phf7* male-specific transcript




**D** Differential expression of *Sxl*'s male-specific exon




**E** Differential expression of C*lic* exon 1



## Discussion

This RNA-Seq experiment was undertaken to find targets of *Sxl* in the germline, with the greater goal of understanding how *Sxl* controls germline sexual identity. This goal was completed successfully as analysis of the sequencing data highlights many genes whose expression changed in response to loss of *Sxl* from the germline. This list of genes represents possibble targets of *Sxl* in the germline (Appendices A-B). In addition, I confirmed that CG15930, one of the hits from the RNA-Seq, is a target of *Sxl* in the germline, and also functions in germline sexual identity (discussed in Chapter 3).

The finding that many uncharacterized genes are upregulated in *Sxl*-RNAi suggests that *Sxl*'s targets in the germline have remained unknown to date because these genes have simply not been studied. This possibility holds up under scrutiny, since the first gene discovered by our lab to have a role in germline sexual identity was a previously uncharacterized gene *CG9576*. This gene now named *Phf7*, promotes male identity in the male germline (Yang et al, 2012). My RNA-Seq data showed that the *Phf7* male-specific transcript (*Phf7*-RC), is repressed by *Sxl* in the female germline (Fig 2.7C). Another previously uncharacterized gene *CG15930*, was discovered in this work as being highly expressed in the *bam* male RNA-Seq compared to females, and is highly upregulated in the *Sxl*-RNAi ovaries compared to control-RNAi. I undertook an

in-depth study of the function of this gene (discussed in chapter 3) and discovered that like *Phf7*, *CG15930* promotes male identity in the male germline, and is repressed by *Sxl* in the female germline.

The parallels between *CG15930* and *Phf7* are striking. They are both previously uncharacterized genes on the X chromosome that promote male identity in the germline, and are both repressed by *Sxl*. These observations suggest that one mechanism by which *Sxl* promotes female identity in the germline is by repressing male-promoting genes (Fig 2.8). This hypothesis is supported by the global analysis of my *Sxl* RNA-Seq data, which shows a higher percentage of genes upregulated in *Sxl*-RNAi ovaries. An important next step for this study is to determine whether any of the genes that are upregulated in *Sxl*-RNAi ovaries are able to promote male sexual identity in the germline.

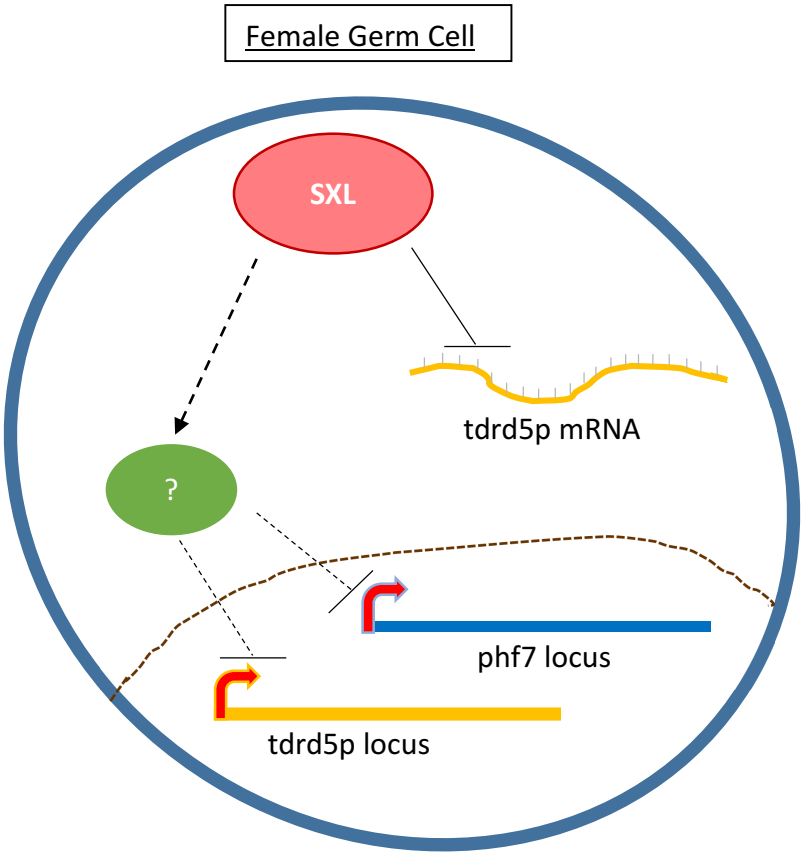



**Fig 2.8: Model of how SXL promotes female identity in the germline.** SXL plays a largely repressive role in the female germline. It acts through an as yet unknown gene to repress expression of genes which promote male identity in the germline. Two such male-promoting genes have already been validated, Phf7, and tdrd5p (discussed in Ch. 3). SXL may also directly repress translation of certain target genes in the germline (discussed in Ch. 3).

# Chapter 3:

# *tudor5-prime* promotes male identity in the germline

## Introduction:

The RNA sequencing experiment discussed in chapter 2 produced many candidate genes whose expression appear to be regulated by SXL. One such gene is a previously uncharacterized gene currently called *CG15930*. *CG15930* is a small (2kb) gene which makes a tudor domain-containing protein. Tudor domain proteins have very important roles in the germline of many species. The founding member of this family of proteins, Drosophila tudor, is essential for germ cell formation (Raff et al, 1990; Breitwieser et al, 1996; Mahowald, 2001). CG15930 has been found as differentially expressed in published large-scale sequencing works in the brain (Janic et al, 1994) as well as in purified germline stem cells (Kai et al 2005), however, no further studies have been reported for it. The upregulation of *CG15930* has also been reported in ovaries mutant for *snf* (Kai et al 2005, and Chau et al, 2009), a general splicing factor that also plays a role in the splicing of *Sxl*.

In my RNA-Seq data, *CG15930* is 16 fold upregulated in the *Sxl* RNAi dataset compared to control RNAi. Its expression in *Sxl* RNAi ovaries is closer to the expression levels of *bam* mutant males. Interestingly, the fold enrichment of *CG15930* in gonads of *bam* mutant males compared to *bam* mutant females is also exactly 16 fold. These data suggest that loss of *Sxl* caused *CG15930* to switch from the female mode of expression to the male mode of expression. In addition,

*CG15930* mRNA has 2 predicted *Sxl* binding sites, one in the 3rd intron and the other in the 3' UTR, suggesting that it's regulation by *Sxl* may be direct. Regardless of whether *CG15930* is a direct or indirect downstream target of *Sxl*, it is an enticing gene to study the role it plays in germline sexual identity. In this chapter, I discuss my investigation into the function of CG15930 in males and its role in promoting male sexual identity in the germline. My analysis of this gene suggests that it does play a role in promoting male identity and development of the male germline, a likely explanation for why it is normally repressed in females. Since this is the first tudor domain protein shown to play a role in sexual identity in the germline, we have decided to name this gene *tudor5-prime (tdrd5p)*.

## Materials and Methods

### Fly stocks and fecundity tests

The fly stocks used were obtained from Bloomington Stock Center unless otherwise indicated. $bam^{1}$ (McKearin & Spradling, 1990), $bam^{\Delta 86}$ (BDSC# 5427), *nos*-Gal4 (Van Doren et al, 1998), the control RNAi used was p{VALIUM20-mCherry}attP2 (BDSC# 35785), uas-*Sxl*-RNAi=TRiP.HMS00609 (BDSC# 34393), uas-CG15930-RNAi=TRiP. GL01046 (BDSC# 36882), $Snf^{148}$, $otu^{17}$, attP40{*nos*-Cas9} (NIG-FLY# CAS-0001).

Fecundity tests were carried out by setting up crosses with one *tudor5-prime* mutant male and 15 virgin females of the control stock. The control stock used is nos-Cas9 isogenized to FM7KrGFP fly stock to replicate the treatment of the *tudor5-prime* mutant fly lines while screening them for transformants (heretofore referred to as nos-Cas9-iso). Fecundity was determined for five *tudor5-prime* mutant males compared to 5 control males. The test was designed to stress the sperm-production capacity of each male. Males were aged with females to allow mating during the ageing period (this is important to minimize amount of stored sperm). After aging, each test male was placed in a bottle with 2 day old virgin females and allowed to mate for 3 days. Females were then discarded and each male was placed with another 15 virgin females in a new bottle. This was repeated two more times for a total of four mating bottles per male. Offspring were counted while still stuck to the wall of the bottle at the black pupa stage. Total offspring per male was calculated by summing the number of offspring from each of the four mating bottles.

Immunofluorescence

Adult ovaries and testes were fixed, blocked and stained as described in Gonczy et al, 1997. All images were taken with a Zeiss LSM 510 confocal microscope. Primary antibodies and the concentrations used are as follows: chicken anti-Vasa 1:10,000 (K. Howard); rabbit anti-Vasa 1:10,000 (R. Lehmann); rat anti-Ncadherin

1:12 (DN-EX#8, DSHB); rat anti-HA 1:100 (3F10, Roche); mouse anti-HTS 1:4 (1B1, DSHB); mouse anti-Armadillo 1:100 (N2 7A1, DSHB). DSHB: Developmental Studies Hybridoma Bank. Secondary antibodies were used at 1:500 (Alexa-fluor). Stains were mounted in vectashield mounting solution with DAPI (vector Industries).

## RT-PCR, quantitative RT-PCR & In-situ hybridization

For RT-PCR and qRT-PCR, total RNA was isolated from *bam* mutant ovaries and testes using RNA-bee (Tel-Test). Contaminating DNA was removed from the RNA using Turbo-DNA-free (Ambion). RNA was converted to cDNA using Superscript II (Invitrogen). qRT-PCR was done using 2 biological replicates and in technical triplicate.

In-situ hybridization was done as described in Morris et al, 2009. DIG-labelled sense and antisense probes were synthesized from PCR products. Primers used for PCR; forward: CCATACGACGATCAGCAGCT, reverse: CTCGACCATCCCAAAAGGCT

## Mutagenesis

Mutant alleles of *tudor5-prime* were created using CRISPR-Cas9 mediated genome editing. Small guide RNA (sgRNA) was designed and cloned following the Perrimon lab protocol (Ren et al, 2013), using the U6b-sgRNA-short vector

described in publication. The sgRNA was injected by Best Gene inc into *nos*-Cas9(II-attP40) flies. The guide RNA was designed against the first exon (Fig 3.9A) to increase the chance of a frame shift caused by a mutation early in the gene. This outcome was achieved (Fig 3.10). Screening to test for mutations in the *tudor5-prime* locus was done using PCR instead of the hgma technique described in the Perrimon lab's paper. The left primer was designed so the 3' end of the primer ended one nucleotide after the expected Cas9 genomic cut site (Fig 3.9B). In this way the 3' end of the primer would not be able to anneal, thereby preventing its extension by polymerase. This screening method worked very well; PCR was unsuccessful for the flies with a mutation in the *tudor5-prime* locus caused by Cas9 nuclease activity (Fig 3.9B).

## BAC-tagging

The *tudor5-prime*:HA transgenic flies were generated by BAC recombineering using the RP98-1M22 BAC obtained from the BACPAC Resources Center (Venken et al, 2009). A 3xHA epitope tag was added to the N-terminus of the gene. This construct was also modified to delete the SXL binding sites in the intron and/or 3'UTR using QuickChange II site-directed mutagenesis kit (agilent). These constructs were all generated by another member of our lab: Kelly Baxter.

## Results

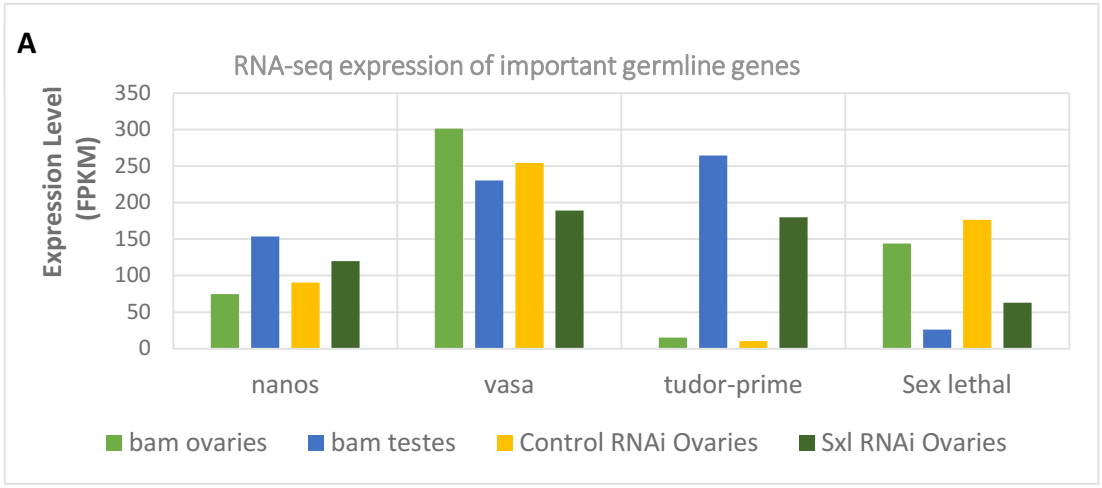
*Tudor5-prime* is expressed in a sexually dimorphic manner

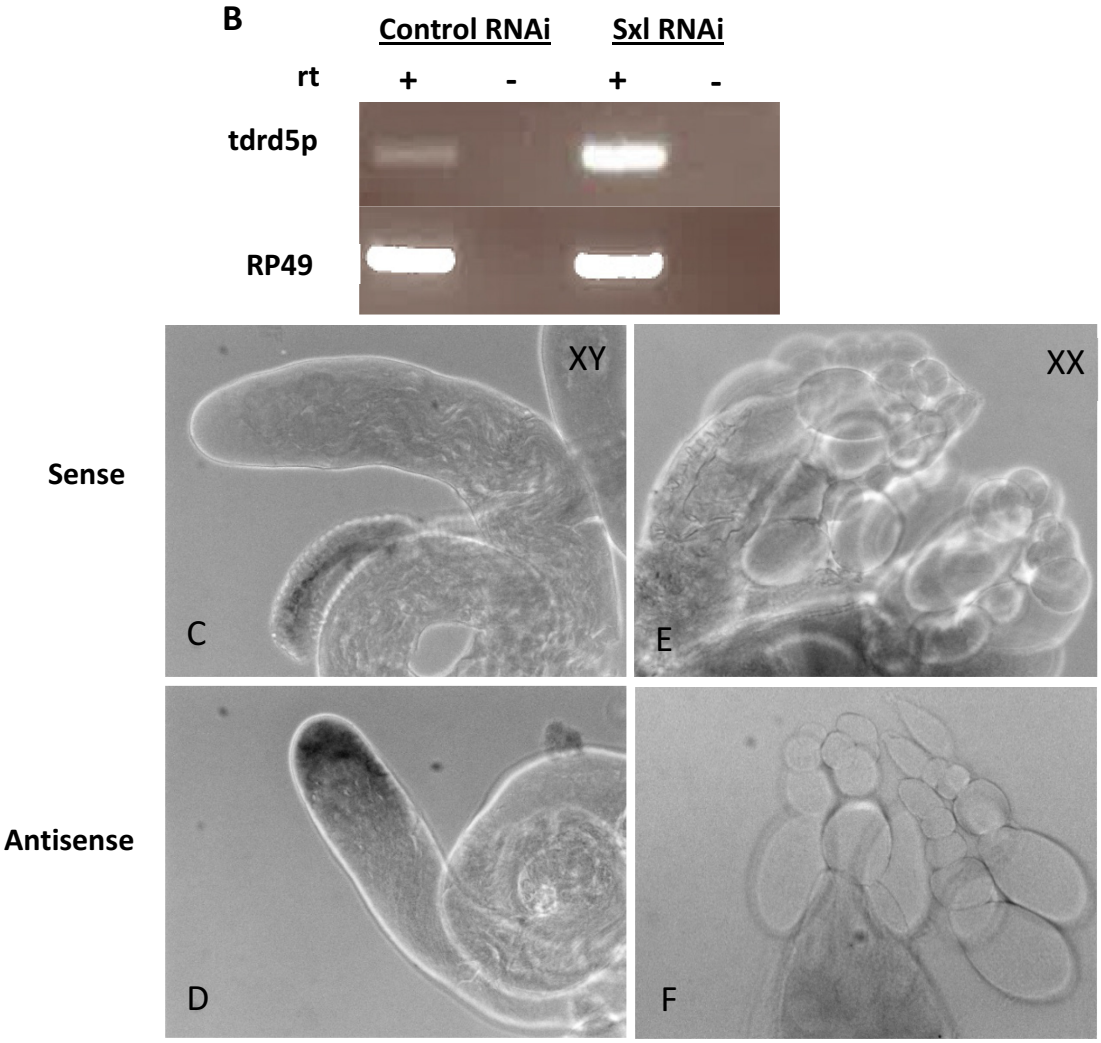
Several pieces of data show that *tudor5-prime* is expressed in a sexually dimorphic manner; it has higher expression in males at the RNA level as well as at the protein level. Its RNA-Seq expression profile shows that it is not merely 16 fold more highly expressed in bam mutant males compared to females, but in fact its overall expression level in males is quite high. It has an FPKM (fragments per kilobase of transcript per million mapped reads) close to 300 in males. Comparison of *tudor5-prime*'s RNA expression to that of other genes known to have important functions in the germline shows that this level of expression is substantial. In *bam* males *vasa* has an FPKM of 250, while *nanos* is expressed at an FPKM of 150 (Fig 3.1A). These two genes are known to have very important functions in the germline. Additionally, the comparison of genes that have important but non sex-specific functions in the germline show a similar expression level in all of the genotypes used for the RNA-Seq experiment. This confirms that these genes are important for proper germline development in both sexes. The expression of *tudor5-prime*, however, is decidedly more sex-specific—it is highly expressed in *bam* testes compared to *bam* ovaries, and is upregulated in *Sxl*-RNAi ovaries (Fig 3.1A). This sex-specific expression pattern suggests that

*tudor5-prime* is needed male-specifically. It also suggest that this male-specific gene is normally repressed by SXL in the female germline.

The sex-specific expression pattern predicted by the sequencing data was validated by RT-PCR and in-situ hybridization (Fig 3.1 B, C). The RT-PCR was done using the same two genotypes used for the *Sxl*-RNAi RNA-Seq experiment: both genotypes are in the *bam* mutant background; the control genotype expresses a control RNAi in the germline, and the experimental genotype expresses Sxl RNAi in the germline. The RT-PCR shows *tudor5-prime* upregulated in ovaries lacking *Sxl* function in the germline. (Fig 3.1B). The in-situ hybridization was done on wild type gonads. Since the RNA-Seq experiment was done solely in the bam mutant background, it was important to assess the expression of *tudor5-prime* in a wild type background. To accomplish this, I performed an in-situ hybridization experiment on wild type testes and ovaries. The in-situ showed that similar to *bam* mutant gonads, *tudor5-prime* is expressed male-specifically in wild type gonads. (Fig 3.1C-F).

**Figure 3.1: *tudor5-prime* is expressed more highly in males at the RNA level**. A) The expression profiles of genes with important germline functions. B) Validation of the *tdrd5p* RNA-Seq expression data by RT-PCR. Genotypes used are *bam*-/-, nosGal4> mCherry control RNA and *bam*$^{-/-}$, *nos*-Gal4>*Sxl* RNAi. C-F) In-situ hybridization showing expression of *tdrd5p* mRNA in C-D)wild type testes and E-F) wild type ovaries.


A
RNA-seq expression of important germline genes
Expression Level (FPKM)
350
300
250
200
150
100
50
0
nanos
vasa
tudor-prime
Sex lethal
bam ovaries
bam testes
Control RNAi Ovaries
Sxl RNAi Ovaries


B
Control RNAi
Sxl RNAi
rt
+
-
+
-
tdrd5p
RP49
Sense
Antisense
XY
XX
C
E
D
F

To determine the expression pattern of *tudor5-prime* at the protein level, I analyzed the expression of a TDRD5P:HA fusion protein made by BAC recombineering (by Kelly Baxter). This TDRD5P:HA fusion protein shows that TDRD5P protein is expressed in the germline of both males and females, but at a much higher level in males (Fig 3.2). Since this is a genomic construct, it should recapitulate the endogenous expression pattern of the TDRD5P protein. In males, the protein is expressed at a slightly lower level in germline stem cells, and a higher level from the 4 cell cyst stage and beyond. The protein accumulates into distinct foci that are predominantly cytoplasmic with some abutting a perinuclear germline structure called the nuage. There are also a few foci located within germ cell nuclei. Unlike wild type ovaries, TDRD5P protein is very highly expressed in ovaries that lack Sxl function in the germline (Fig 3.3). This data shows that the sex-specific RNA expression results in sex-specific protein expression as well. The expression pattern of *tudor5-prime* suggests that it is regulated by *Sxl* in the female germline.

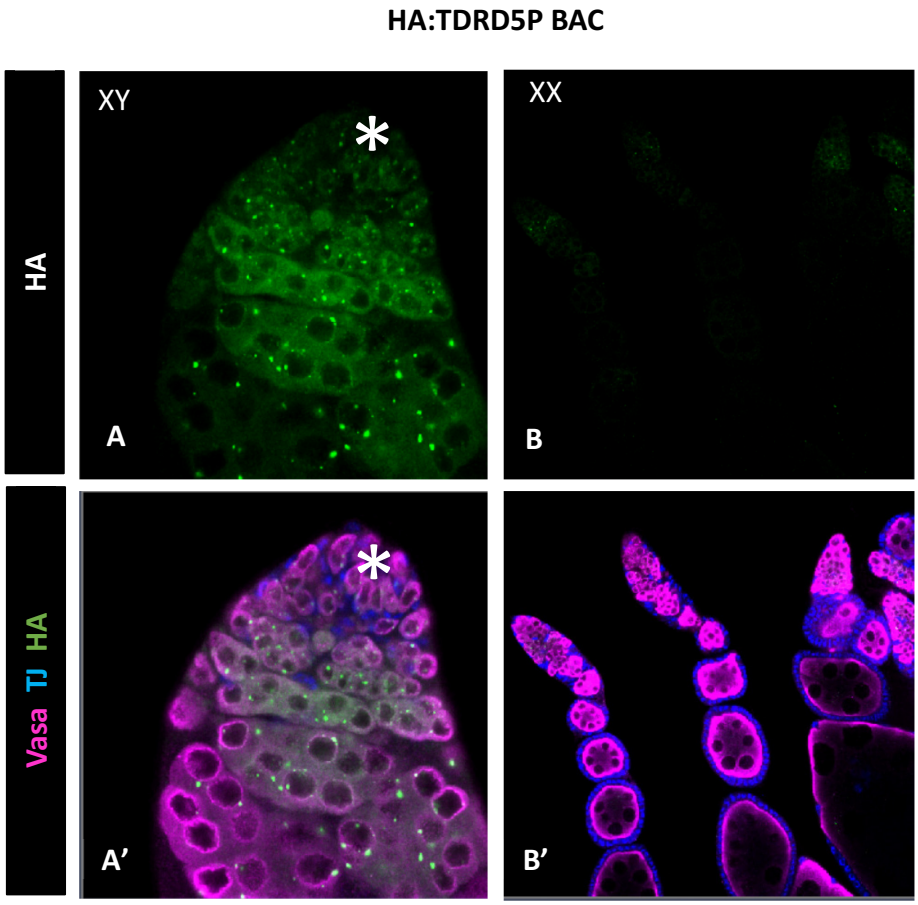



**Figure 3.2: Tudor5-prime is expressed more highly in males at the protein level**. Expression of the HA tagged TDRD5P protein in A) wild type testes and B) wild type ovaries. HA stains TDRD5P, Vasa stains germ cells, TJ stains somatic cells.

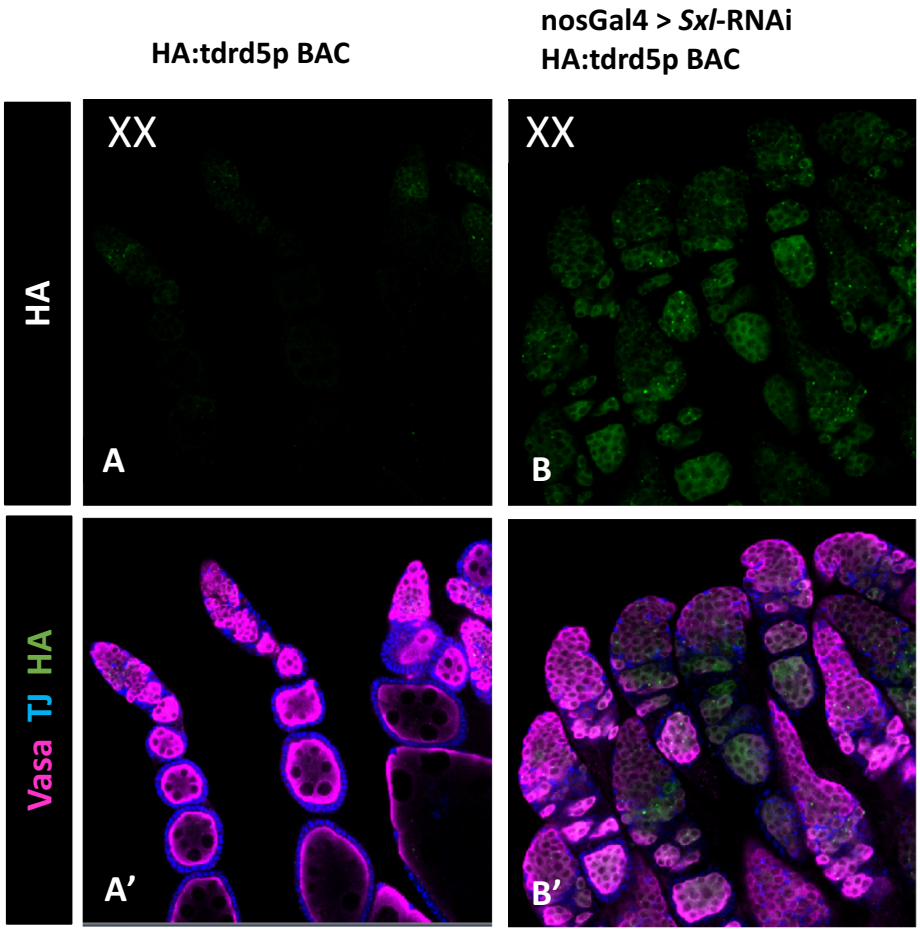



**Figure 3.3: *Tudor5-prime* expression is regulated by *Sxl* function in the female germline**. Expression of the HA tagged TDRD5P protein in A) wild type ovaries and in B) ovaries with *Sxl* knocked down in the germline. HA stains TDRD5P, Vasa stains germ cells, TJ stains somatic cells.

To test whether *tudor5-prime* can be similarly regulated by *Sxl* in male germ cells, I ectopically expressed *Sxl* in the male germline and stained for expression of the TDRD5P:HA construct. This resulted in a dramatic downregulation of TDRD5P:HA fusion protein, specifically in the germline stem cells and early germline where SXL is expressed (Figure 3.4 C-D). This shows that ectopic expression of *Sxl* in the male germline is sufficient to downregulate TDRD5P:HA protein expression. This is also very strong evidence supporting that tdrd5p is a target of Sxl. But is it a direct target of *Sxl*?

From the RNA-Seq data, it is unclear whether *Sxl*'s regulation of *tudor5-prime* is direct or indirect. SXL has not been shown to have the ability to affect transcript abundance directly; it has only been shown to affect splicing and translation. *Tudor5-prime* has two predicted SXL binding sites: one in the 3rd intron and the other in the 3'UTR (Fig 3.4A). This suggests that SXL may be able to bind directly to *tudor5-prime* mRNA. If so, one way that SXL can affect transcript abundance of *tudor5-prime* is to cause degradation of the *tudor5-prime* mRNA. Alternatively, SXL may bind and prevent translation of *tudor5-prime*. The downregulation of TDRD5P:HA upon the ectopic expression of *Sxl* in male germ cells could result either from direct translational regulation or from indirect transcriptional regulation. The most straightforward way to answer this question is to delete the SXL binding sites and test for loss of regulation.

While these deletion mutants are being generated, another way to test this hypothesis is by RT-PCR of *tdrd5p* RNA. I performed an RT-PCR experiment on testes with ectopic expression of *Sxl* in the germline. The RT-PCR showed that *tudor5-prime* mRNA is unchanged in testes expressing *Sxl* compared to control testes (Fig 3.4B). This data suggests that *Sxl* is unable to regulate *tudor5-prime* RNA expression in a male environment. This is possibly because SXL's transcriptional regulation of *tudor5-prime* requires the activity of other genes that can only be found in the female germline. An alternative hypothesis is that the restricted area of repression caused by ectopic *Sxl* expression, is too small to see a change by RNA (see fig 3.4D). Therefore, the level at which *Sxl* regulates *tdrd5p* is still unknown.

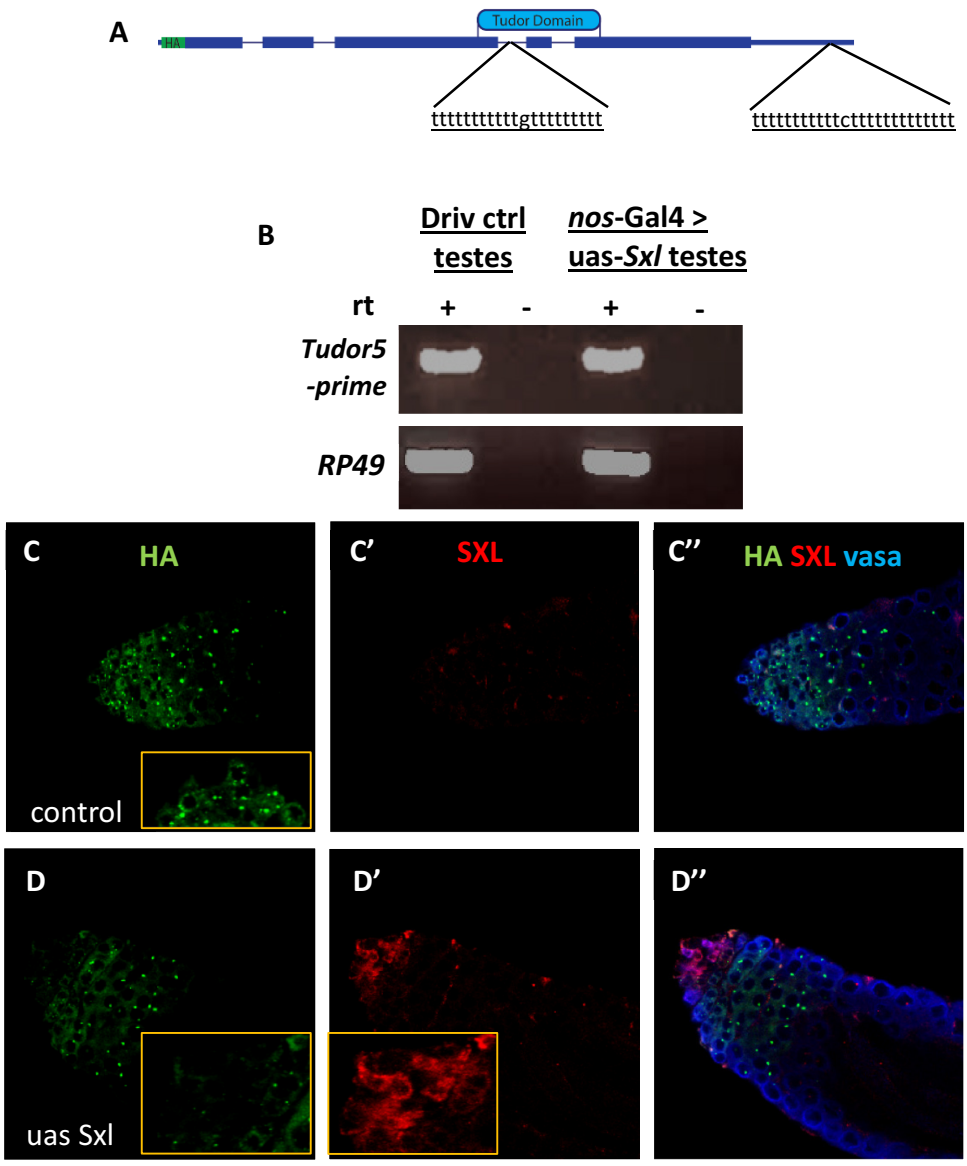



**Figure 3.4: *tudor5-prime* expression is regulated by Sxl in the male germline** A) Location of two putative Sxl binding sites in the *tudor5-prime* transcript. B) RT-PCR of *tudor5-prime* in *nos*-Gal4 driver only control testes compared to testes with ectopic expression of *Sxl* in the germline. C-D) Confocal images of C) *nos*-Gal4 HA:*tdrd5p* control testes and D) *nos*-Gal4 > uas *Sxl*; HA:BAC testes. HA stains TDRD5P, Vasa stains germ cells.

Ectopic expression of *tudor5-prime* in females

SXL strongly represses the expression of *tudor5-prime* in the female germline. This suggests that the expression of *tdrd5p* may be detrimental for female germline development. To investigate this possibility, I generated a transgenic fly expressing genomic *tdrd5p* under UAS control. When *tdrd5p* is ectopically expressed in the female germline using a *nos*-Gal4 driver, no morphologic defect is observed. Interestingly, when *tdrd5p* is ectopically expressed in the female soma using *traffic-jam* (tj) Gal4, gross morphological defects are observed. Late stage egg chambers die or do not develop normally in these ovaries (compare Fig 3.5 A to B, arrows), so these females are sterile. The degenerating egg chambers are clearly seen by comparing the Dapi channels (Fig 3.5 A' and B'). The control ovaries show well-formed egg chambers, while the overexpression shows many tiny blebs of Dapi, representing dying cells. As these females age, germline loss progresses dramatically, so that by 7 days old none of the tj>tdrd5p ovaries look completely wild type. These defects suggest that the soma has low tolerance for the expression of *tdrd5p*, and is likely linked to the as yet undetermined function of *tudor5-prime.*

Additionally, in *tj*-gal4>*tdrd5p* females there is a significant upregulation in the expression of pMAD—a downstream effector for BMP signaling—in the soma as well as in the germline of late stage egg chambers (Compare Fig 3.6 A' to

B', arrowheads). This suggests transcriptional upregulation of the BMP pathway, as well as soma–germline signaling in response to the ectopic expression of *tudor5-prime* in the soma. Intriguingly, pMAD expression in the germarium, retains its normal expression pattern in most *tj*-gal4>*tdrd5p* females; it is restricted to germline stem cells GSCs (Compare Fig 3.6 A' to B', arrows). pMAD is also restricted to GSCs in the germaria of ovaries lacking *Sxl* in the germline (Fig 3.6C', arrows). This suggests that regulation of the BMP signaling pathway is upstream of or parallel to *Sxl* function in the germline. The TRIM-NHL protein *Brain tumor* (*Brat*) has been shown to antagonize bmp signaling in cystoblasts by repressing translation of the bmp effector *Madea* (Harris et al, 2011, Li et al, 2012). The *fused*/*smurf* E3 ubiquitin ligase complex has also been shown to repress bmp signaling by causing the degradation of the BMP receptor *thickveins* in cystoblsts (Xia et al, 2010). This tight, multilayered regulation of BMP signaling in the early germline is important to maintain the integrity of germline stem cells and allow cystoblasts to differentiate, and it may explain why ectopic expression *tudor5-prime* is not able to upregulate BMP signaling in the early germline.

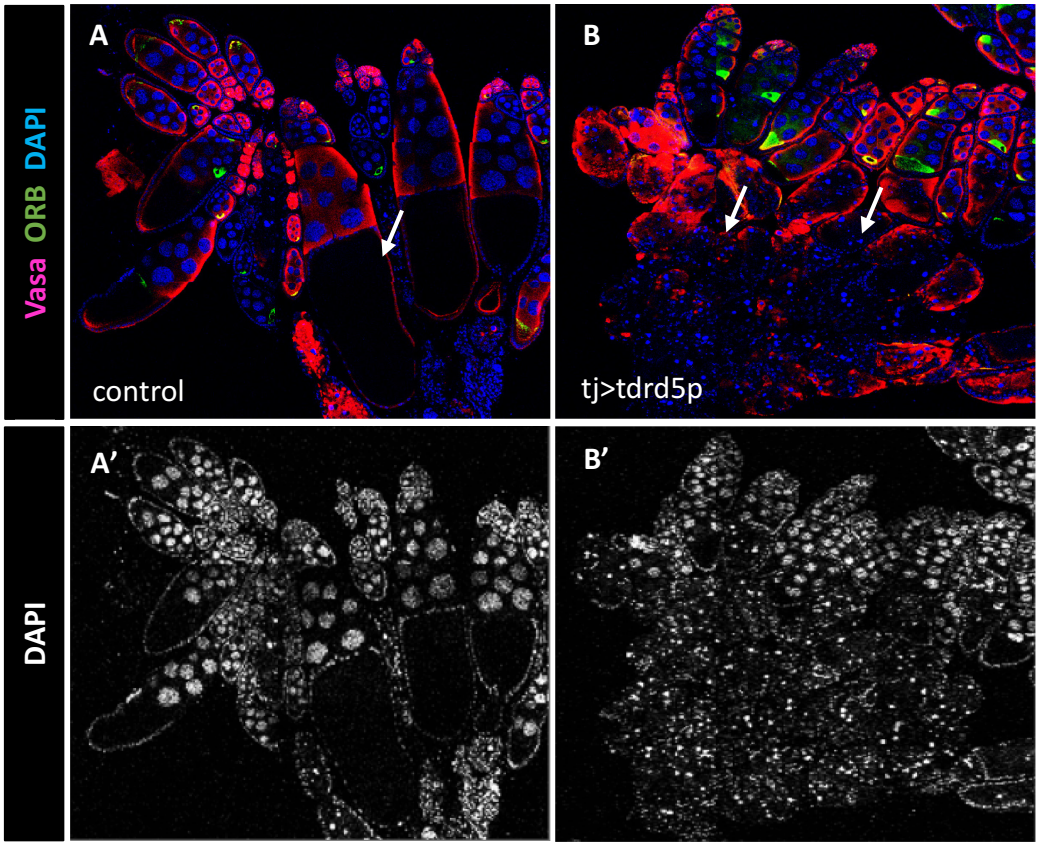



**Figure 3.5: Overexpression of *tdrd5p* in the female soma results in death and sterility** A-B) Confocal images of A) driver only control and B) *tj*-gal4>*tdrd5p* ovaries. Vasa stains the germ cells, Orb stains the oocyte, Dapi stains DNA.

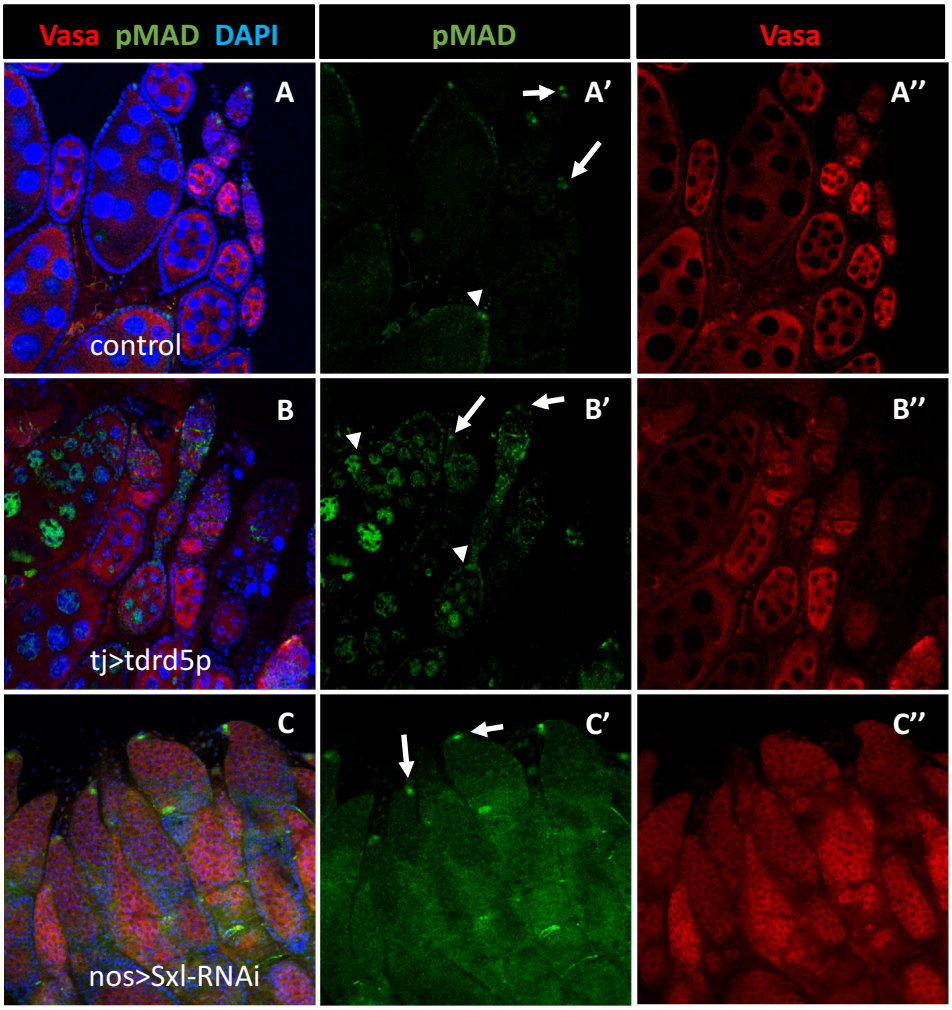



**Figure 3.6: Overexpression of *tdrd5p* in the female soma results in over-activation of BMP signaling pathway** A-C) Confocal images of A) Control ovaries, B) ovaries with ectopic expression of *tdrd5p* in the soma, and C) ovaries with *Sxl* knocked down in the germline. Arrows point to pMAD expression in GSCs, arrowheads point to pMAD expression in egg chambers. Vasa stains germ cells, Dapi stains DNA.

## *Tudor5-prime* promotes male identity in the germline

To further investigate *tudor5-prime's* role in sexual identity we decided to conduct our experiments using the sensitized genetic backgrounds frequently used in the investigation of genes involved in sexual identity. Females mutant for *transformer* (*tra*)—a key player in the somatic sex determination pathway—undergo a transformation of the somatic gonad, so that XX *tra* mutants develop testes instead of ovaries. These testes are highly underdeveloped, however, causing these animals to be sterile (Fig 3.7B). It is believed that the reason why these testes are underdeveloped in XX *tra* mutants, is that the germline has a different sexual identity than the soma—the germ cells still have female identity while the soma now has male identity.

A strong test of a gene's ability to promote male identity in the germline is to overexpress it in an XX *tra* mutant background, and analyze these testes for rescue of the germline development defect. Previous work from our lab has shown that *Phf7*, a gene shown to promote male identity in the germline, causes a 6% rescue of the XX *tra* mutant phenotype (Yang et al, 2012). Ectopic expression of *tudor5-prime* in the germline of XX *tra* mutants results in a 18% rescue of the phenotype (Fig 3.7 C, D). This is truly strong evidence supporting *tudor5-prime* as a gene able to promote male identity in the germline. Interestingly, overexpressing *tdrd5p* in addition to *Phf7* resulted in a higher rescue than *Phf7*

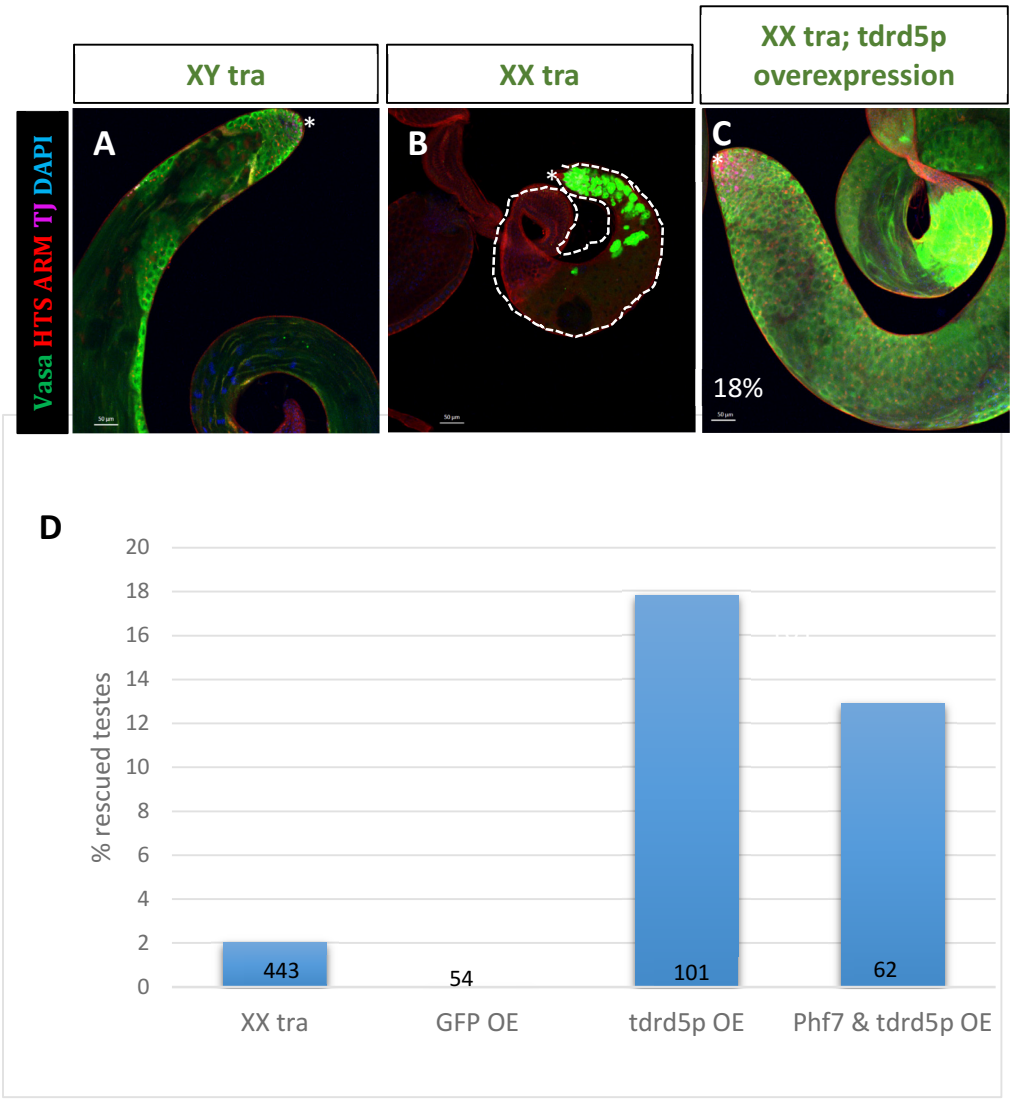


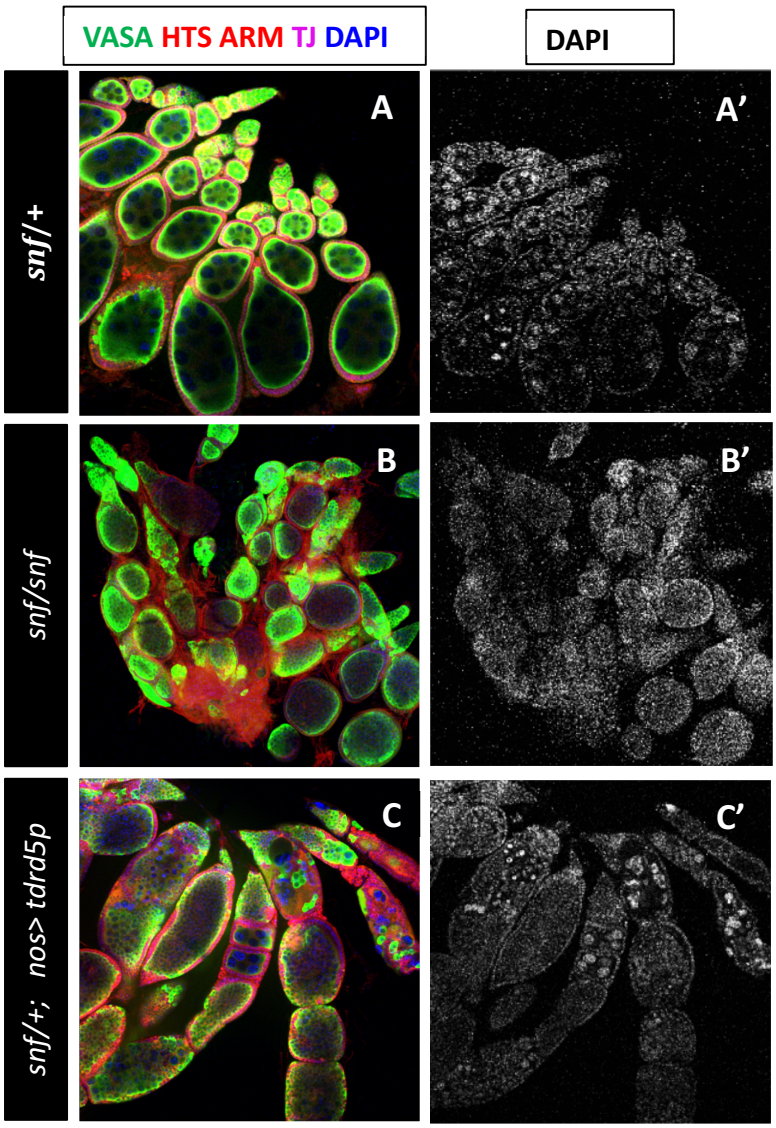

**Figure 3.7: Ectopic expression of *tudor5-prime* promotes male identity in XX *tra* mutants.** A-C) Confocal images of A) wild type XY testes, B) XX *tra* mutant testes, and C) XX *tra* mutant testes rescued by *tdrd5p* overexpression. Asterisk marks the hub. Vasa stains germ cells, HTS and Arm stain membranes, TJ stains somatic cells, Dapi stains DNA. D) Quantification of XX *tra* rescue by *tdrd5p* overexpression, *Phf7* & *tdrd5p* overexpression together, and uas-GFP control overexpression.

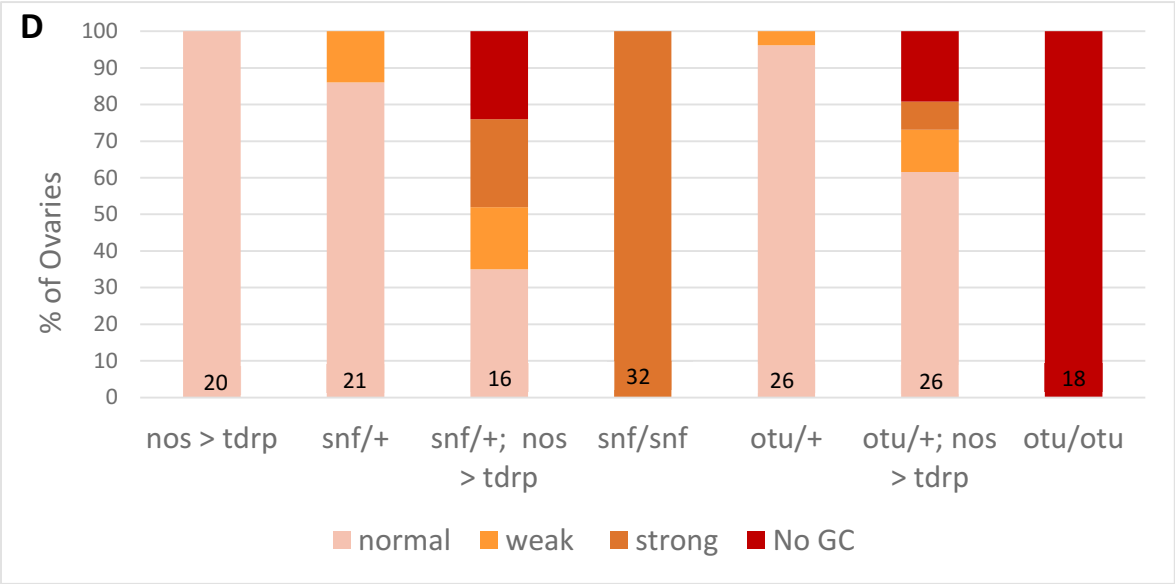
alone, 12%, though lower than *tdrd5p* alone (Fig 3.7D). This intermediate rescue with loss of both male-promoting genes could be caused by a dilution of the GAL4 element to activate 2 UAS elements instead of just one. There also seemed to be a spontaneous suppressor mutation in the *tra* mutant stock which resulted in a 2-3% rescue of the XX tra mutant phenotype. This suppression was lost when uas-GFP was crossed into the *tra* mutant background to use as a UAS control for uas-*tdrd5p* overexpression (Fig 3.7D). It may be useful to map this mutation to determine what gene it is.

Another important test of involvement in germline sexual identity is to determine whether *tudor5-prime* genetically interacts with genes that promote female identity in the germline, *Sxl, otu* and *snf*. Homozygous mutants of these genes have all been shown to cause germline tumors in females (Fig 3.8B). This phenotype has also been associated with germline–soma sexual incompatibility. Heterozygotes of these genes show normal ovary morphology. Ectopic expression of *tdrd5p* in the germline of females heterozygous for *snf*, *otu* and $Sxl^{F4}$ (a germline specific allele of Sxl) results in a dramatic enhancement of the tumor phenotype. 25% of snf/+; nos>*tdrd5p* ovaries have large pervasive germline tumors similar to the homozygous mutants (Compare Fig 3.8 B and C = tumorous egg chambers to 3.8B = normal egg chamber development). Another 25% of these ovaries show a complete loss of the germline; this phenotype

mimics a strong *otu* mutant phenotype (King et al, 1986, Storto and King 1987). This evidence lends further support that *tdrd5p* promotes male identity in the germline.

**Fig 3.8: Over-expression of *tudor5-prime* enhances the heterozygous phenotype of key germline sex regulators.** A-C) Confocal images of A) *snf*/+ ovaries, B) *snf*/*snf* mutant ovaries, C) *snf*/+ ovaries with ectopic expression of *tdrd5p*, D) Quantification of the degree of enhancement of the phenotype to tumorous ovaries and germline loss in *snf*/+ and *otu*/+ ovaries


VASA HTS ARM TJ DAPI
DAPI
snf/+
snf/snf
snf/+; nos> tdrd5p
A
A'
B
B'
C
C'

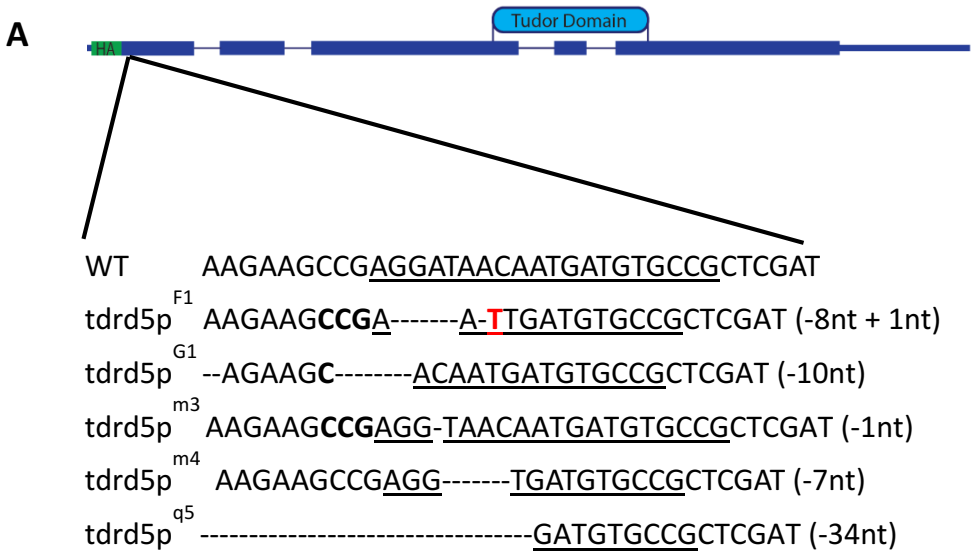

D
% of Ovaries
100
90
80
70
60
50
40
30
20
10
0
20
21
16
32
26
26
18
nos > tdrp
snf/+
snf/+; nos > tdrp
snf/snf
otu/+
otu/+; nos > tdrp
otu/otu
normal
weak
strong
No GC

## *Tudor5-prime* is required for proper male fertility & germline differentiation

The male-specific expression pattern of *tudor5-prime* suggests it has an important function in the male germline. Knocking down *tudor5-prime* function by RNAi did not produce a germline phenotype. RNAi efficacy can be variable and there is usually an element of uncertainty with experiments involving it— mutant alleles are always more dependable. There were no *tudor5-prime* mutant alleles available, however, so we decided to generate *tudor5-prime* mutants using CRISPR-Cas9 genome-editing technique. The *tudor5-prime* genome editing was very efficient— it produced a 70% mutation rate. Using this technique, I generated 28 unique alleles with many very different lesions; a full list of which can be found in table 3.1.

Using several different *tudor5-prime* mutants (Fig 3.9A), I analyzed male fertility and testis morphology of both young males and aged males. Immunohistochemistry on the testes of young males does not show a morphological defect. Older males, however, show several low penetrance defects: 15% of *tudor5-prime* mutant males 15-20 days old have a displaced hub phenotype (Fig 3.11D), 40% show an expansion of the dapi bright region and 7% show a more pronounced expansion of the mitotic germ cells (the dapi bright region) with accompanying germline loss that results in a *skinny testis*

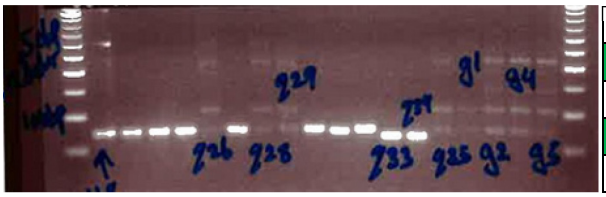
appearance (Fig 3.11B-C). At $29^0$C, *tudor5-prime* mutant testis phenotypes are more severe; almost 20% of testes show severe germline loss, with 13% of these showing complete loss all germ cells (Fig 3.12B-C). These combined phenotypes suggest a defect in proper germline differentiation as well as germline maintenance. In addition to the morphological defects, young (5 days old) *tudor5-prime* mutant males have a 50% reduction in fecundity compared to control (Fig 3.13). This suggests that while the morphological defects present at a low penetrance, their overall effects still culminate into a substantial reduction in fertility; a phenotype which supports *tudor5-prime*'s importance for male germline development. RNAi-mediated knockdown of *Phf7* (another gene recently discovered by our lab to be important for male-specific germline development), did not enhance the *tdrd5p* phenotype in 7 day old flies.




**B**

Genome:AAGAAGCCGAGG|ATAACAATGATGTGCCGCTCGAT

Primer: AATTTCGGAAGAAGCCGAGGA




**C**

| # of unique alleles | 28 |
|---|---|
| Smallest deletion | 1nt |
| Largest deletion | 34nt |
| Smallest insertion | 1nt with a deletion |
| Largest insertion | 27nt with 5nt deletion |

**Figure 3.9: *tudor5-prime* mutants used in this study.** A) Sequence of the *tdrd5p* mutants analyzed during my investigation into *tudor5-prime* function. Dashed lines = deleted nts, red letters = inserted nts. A full list of mutants generated can be found in table 3.1. The underlined portion is the genomic sequence used for the sgRNA. B) PCR screening strategy used to detect mutants. Mutants q33 and q34 on the gel are missing 9nt from the cut site. As the 10$^{th}$ nt is also an *A*, these PCRs still worked, producing a PCR product 9nt shorter than expected for a wild type sequence. C) Summary of the *tdrd5p* mutant alleles created using CRISPR-Cas9 genome editing.

**Figure 3.10: Categories of *tudor5-prime* mutants.** Category #1 mutants (see table 3.1) carry a deletion of 3 or more nucleotides without a downstream frameshift. Category #2 and #3 mutants cause a deletion of 1 or more nucleotides that result in a shift to the second or third frame for protein translation. These 2 frames code for premature stop codons throughout the transcript beginning in the first exon. Asterisks (*) represent stop codons.

>Wild Type Protein Sequence
MQSKMDVKEGRLQTLLSKLRKISEEAEDNNDVPLDLRCPSAKM
RNNKVLADLSDMLKVVEDIALEEKTTDSKKPANSEQQINGTIEL
ALDLRQPNAKIRLVISDSSDLQKVFEKAHIALELKTTDSKPTAET
DDQHHVRIFDKMLCNIVVPPDVADVQVNGKPLEQLLATVSERQ
AMNNPGPMKMHELHTGNADFMNSNSQVAEAAASVETLLASTS
DNDSDSDGLESKTTDGYIVKELIPSFSDFPDIGDVSLCALMYGLP
MDAVGMHWKLPPQRIQDICKEDSIFPIIMSCVFSPCEFWFHIVPP
QYAKNPVAEMTIDLNWFYRHTTISSYRAELPSYFYKEGYICAAYS
ECGWRRAMVLVTAPLDAQCVNIEYVDHALSVTLAPNHLRFLPL
SFARTPPLVFRGKMSHVRPLANGWPKNGITAFQRMTFNRVLYA
HVGELDTAQGIFSMRLSYDETFVPTINDLIESRIKLESCCYAPELQ
PFGMVEPLIMPLHDYDDDDDDDDDFAFVP

>Translation for frame change #2-MCRS
MQSKMDVKEGRLQTLLSKLRKISEEAEITMMCRSICAAPVPKCGI
TKC*PI*VTC*KLSKISLWRKKLPIPRSRLILNSRSMEQLN*HSICA
SPMPKYGL*FPIRVTCKKFSKKHISLSS*KQPIPSRRLRLTTSTMC
EFSIRCFATSLCRRM*PMSR*MESHWSSC*PPCRSARQ*TILAQ*R
CMNCTPAMLTS*IAIAKWQKLLRQWKHC*RRHRTTIATAMDSN
QKRRMVTLSRN*YRRFPISRISVMLVCVH*CMDCPWMPLACTGS
CRRSVSRISARRTLSSRSS*AVCSLRVSSGSTLCHRSTPRIL*PR*PL
I*TGSIAIRRSAATVPSCLAISIKRATYAPPTASVAGDVPWCSSPHR
WTPNV*T*SMWIMPCLSHWHRTIYDSSH*ALRAHRHWYFAAK
CRTYAHWPMGGQKMVSPHSSA*HSIGYCMPMSANWIPLRVYSL
CGYRMMKLLCPPSTI*SSRGSSWSPVAMRPNCSLLGWSSR*LCRC
TIMMMMMMMMMISHSFHX

>Translation for frame change #3-CAAR
MQSKMDVKEGRLQTLLSKLRKISEECAARFALPQCQNAE*QSVS
RFE*HAKSCRRYRFGGKNYRFQEAG*F*TADQWNN*ISTRFAPA
QCQNTACDFRFE*LAKSFRKSTYRSRVKNNRFQADG*D*RPAPC
ANFR*DALQHRCAAGCSRCPGEWKAIGAVVSHRVGAPGNEQSW
PNEDA*TAHRQC*LHE*Q*PSGRSCCVSGNTASVDIGQR*RQRWT
RIKNDGWLHCQGTDTVVFRFPGYR*C*FVCTDVWTAHGCRWH
ALEAAAAAYPGYLQGGLYLPDHHELCVLSV*VLVPHCATAVRQE
SCSRDDH*SKLVLSPYDDQQLPCRVA*LFL*RGLHMRRLQRVWL
ATCHGARHRTVGRPMCEHRVCGSCPVCHIGTEPSTIPPTKLCAH
TAIGISRQNVARTPIGQWVAKKWYHRIPAHDIQ*GIVCPCRRTGY
RSGYILYAAIV**NFCAHHQRSNRVADQVGVLLLCARTAAFWDG
RAANYAAARL*********FRIRSIX

Category 2 and 3 mutants cause premature stop codons within the first exon

These should create truncated, nonfunctional proteins

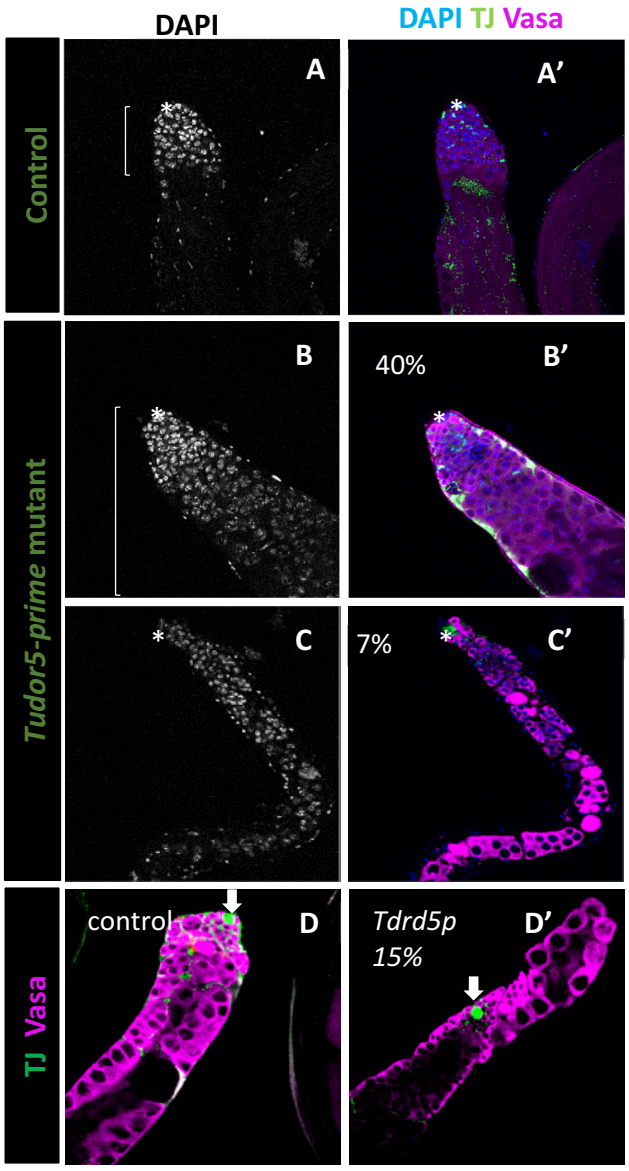



**Figure 3.11: Aged *tudor5-prime* mutant males have differentiation defects**. A-C) 40% of *tudor5-prime* mutant testes show an expanded dapi bright region. 7% show an exacerbated form of this, with some accompanying germline loss. Asterisk marks the hub. D) 15% of *tudor5-prime* mutants show a displaced hub phenotype: n=29. Mutants used: tdrd5p$^{Q5}$ & tdrd5p$^{G1}$. Control genotype: nos-cas9-iso. Arrows point to the hub. Vasa marks germ cells. TJ marks somatic cells.

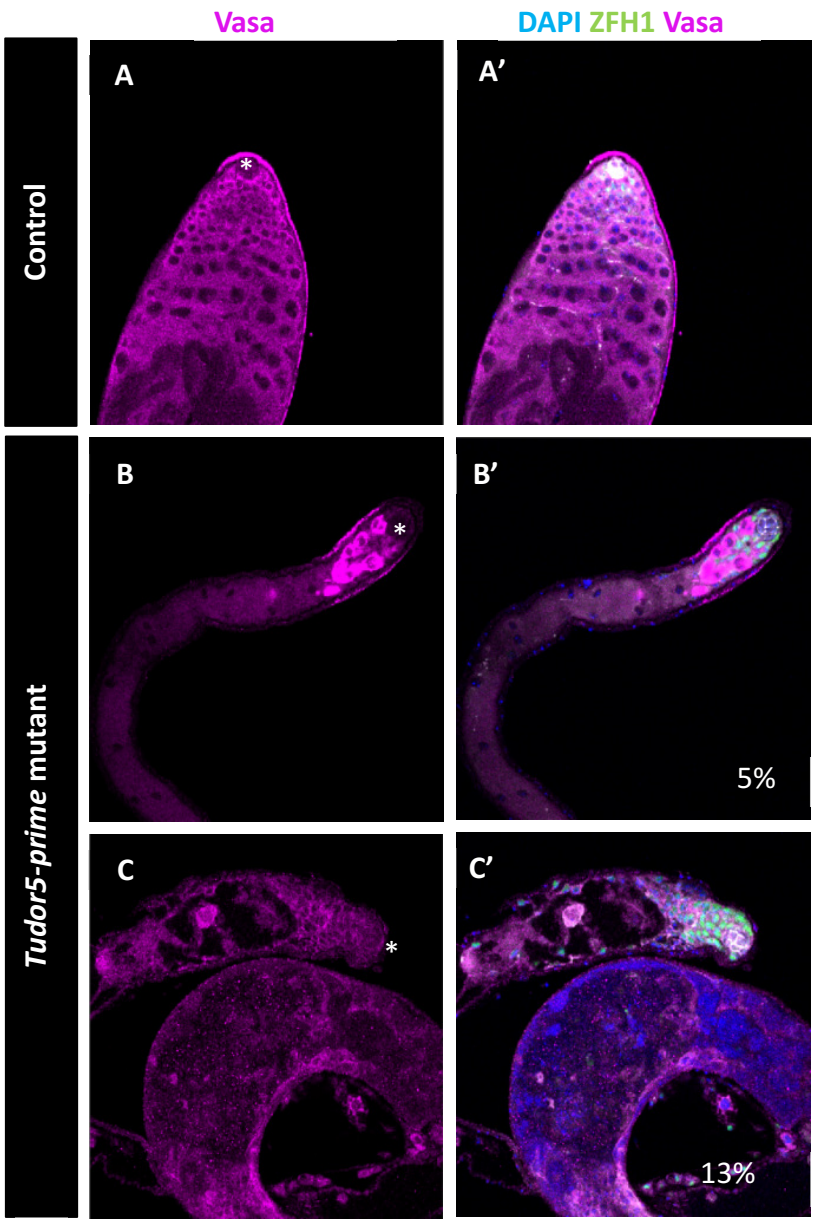



**Figure 3.12: Aged *tudor5-prime* mutant males raised at 29°C have germline defects**. Aged *tudor5-prime* mutant flies raised at 29°C have severe germline defects including complete loss of germline. n=39 testes. Mutant alleles used: tdrd5p$^{F1}$, and tdrd5p$^{m3}$. Asterisk marks the hub. Vasa marks the hub, ZFH1 marks somatic cells.

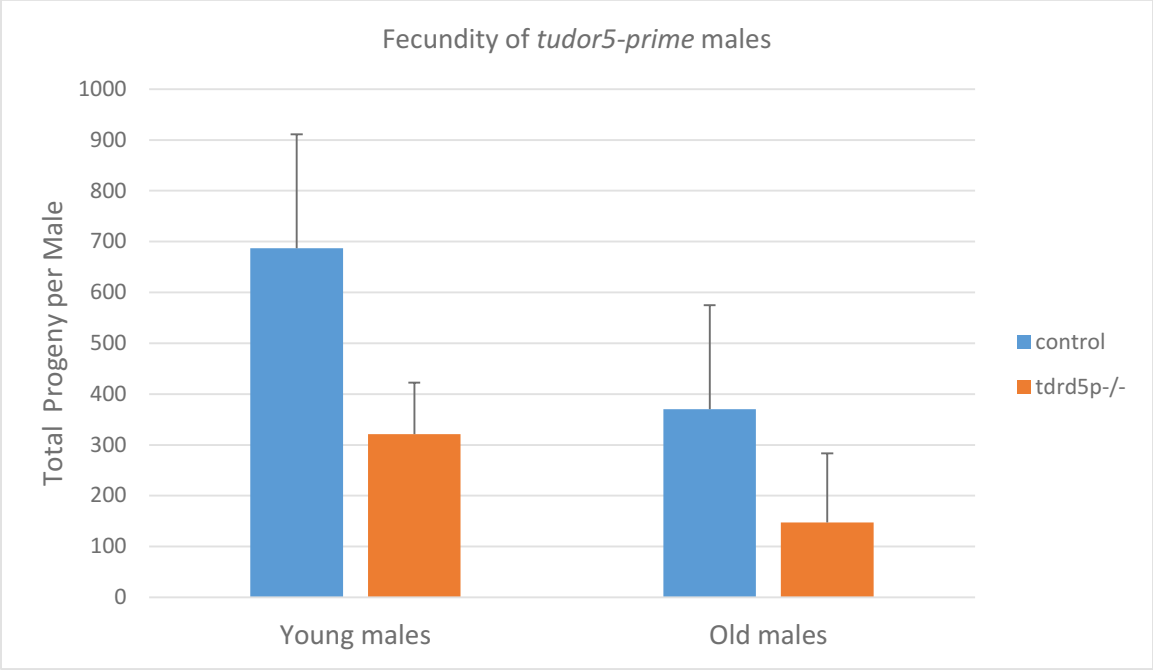
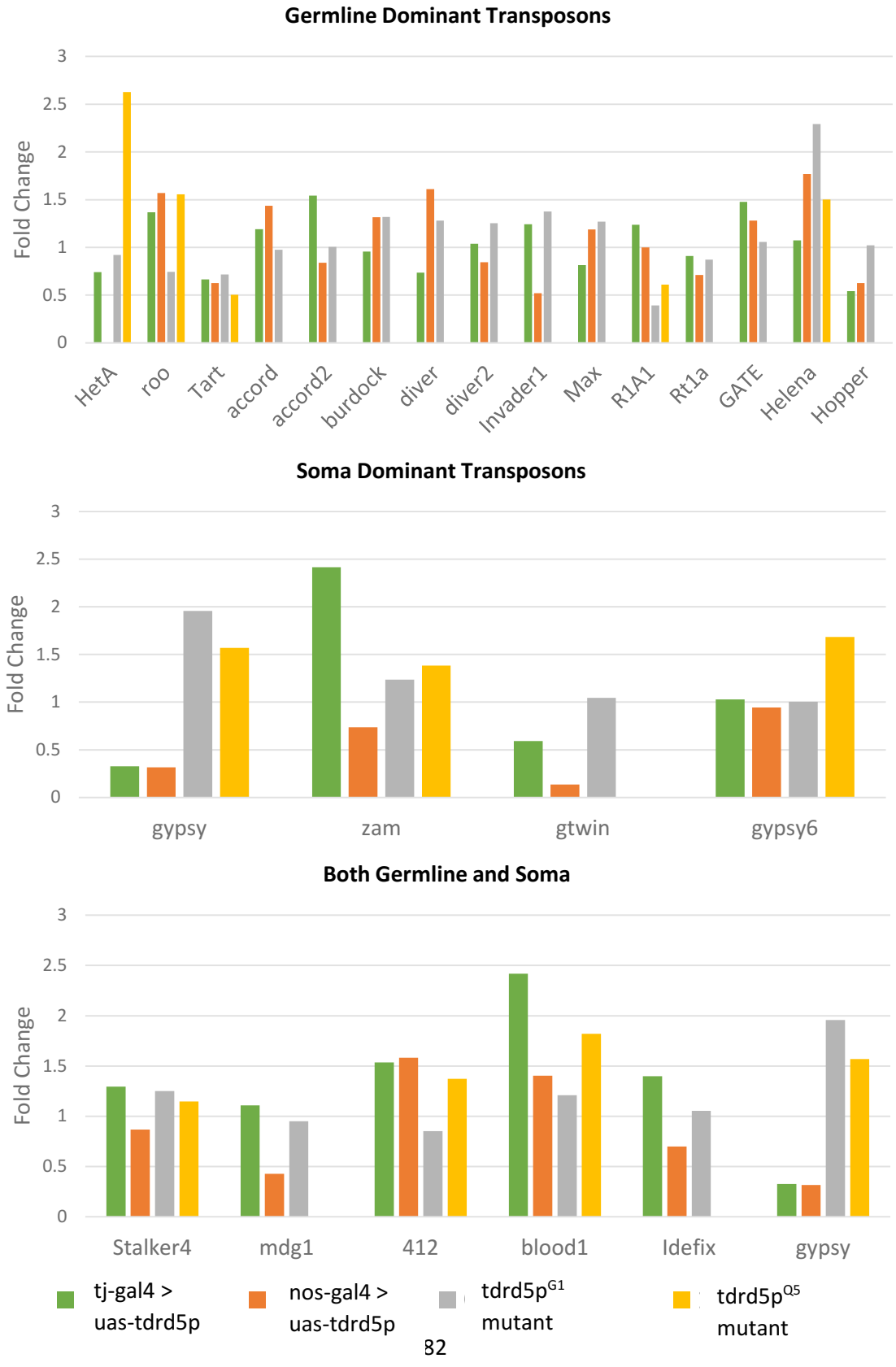



**Figure 3.13: *tudor5-prime* mutant males have low fecundity.** A) *tdrd5p* mutant males have a 50% reduction in fecundity. Young males were 5 days old at the beginning of the fecundity test, aged males were 15 days old. Mutant allele used: tdrd5p$^{m4}$. Control = isogenized *nos*-cas9 males.

Investigating the function of Tudor5-prime

Tudor domain-containing proteins have well known functions in small RNA pathways, transcriptional regulation, and the assembly of snRNPs (Reviewed in Pek et al, 2012). The closest Drosophila gene to *tudor5-prime* is a gene called *tejas*; and the closest mammalian homolog to *tudor5-prime* is mouse TDRD5. Both *tejas* and TDRD5 have been shown to function in the piRNA pathway (Patil & Kai, 2010, Smith et al, 2004, Yabuta et al, 2011). I conducted several different experiments to investigate whether *tudor5-prime's* molecular function was also required in the piRNA pathway. The prime purpose of the piRNA pathway is to repress transposon expression in the germline. Therefore, one common way to determine whether a gene functions in the piRNA pathway is to perform quantitative RT-PCR to measure de-repression of transposon RNA. Transposon expression was not affected in *tdrd5p* mutant males (Fig 3.14). qRT-PCR was also done on ovaries with ectopic expression of *tudor5-prime* either in the germline or in the soma. Neither of these showed any significant change (neither increased nor decreased) in transposon expression.

**Figure 3.14: *tudor5-prime* does not affect transposon expression.** Levels of expression of many different transposons remain unaffected in *tdrd5p* mutant testes, and are also unaffected in ovaries with ectopic expression of *tdrd5p* in either the germline or the soma. Genotypes: *tdrd5p* overexpression in ovaries: *TJ*-Gal4>*tdrd5p*, *nos*-Gal4>*tdrd5p*; *tdrd5p* mutant alleles used: tdrd5p$^{G1}$ & tdrd5p$^{Q5}$. Ovary data compared to Gal4 only driver control ovaries. Testis data compared to FM7KrGFP control testes.


Germline Dominant Transposons
Fold Change
3
2.5
2
1.5
1
0.5
0
HetA
roo
Tart
accord
accord2
burdock
diver
diver2
Invader1
Max
R1A1
Rt1a
GATE
Helena
Hopper
Soma Dominant Transposons
Fold Change
gypsy
zam
gtwin
gypsy6
Both Germline and Soma
Fold Change
Stalker4
mdg1
412
blood1
Idefix
gypsy
tj-gal4 > uas-tdrd5p
nos-gal4 > uas-tdrd5p
tdrd5p$^{G1}$ mutant
tdrd5p$^{Q5}$ mutant

In addition to measuring RNA levels, certain transposon proteins can also be analyzed by immunofluorescence. One such transposon is Stellate; a male-specific transposon known to form crystals in spermatocytes, and it causes male sterility (Bozzetti et al, 1995, Aravin et al, 2001). I examined the formation of Stellate crystals in *tdrd5p* mutant males, as well as *tdrd5p* mutant males also missing one allele of either *tejas*, or *ago3* or *aubergine* (these are two PIWI proteins with key functions in the piRNA pathway). None of these testes showed the formation of Stellate crystals compared to positive control *tejas* homozygous mutant testes, which do show stellate crystals (Fig 3.15). Therefore, while *tudor5-prime's* activity is important for the proper development of the male germline, its specific molecular function remains unknown.

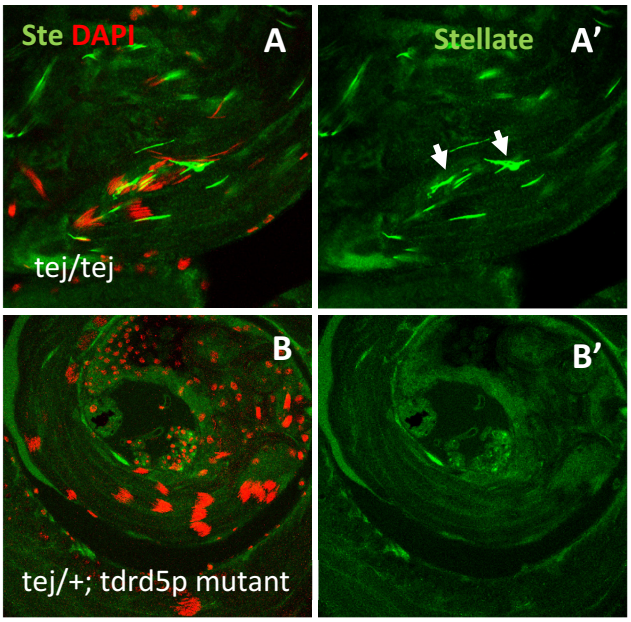



**Figure 3.15: Stellate crystals do not form in *tudor5-prime* mutant testes.** A-B) Confocal images showing stellate crystals in A) *tejas* homozygous mutant testes, but not in B) *tejas* heterozygous testes also mutant for *tudor5-prime*. Arrows mark stellate crystals. Dapi stains condensed sperm heads. Mutant allele used: tdrd5p$^{Q5}$.

## Discussion

In this chapter I showed that *tudor5-prime* plays a role in germline sexual identity. Specifically, it promotes male identity in the germline as overexpressing it can rescue the XX *tra* mutant phenotype, prompting XX germ cells to behave like males. It also enhances the heterozygous phenotypes of genes that promote female identity in the germline. It is likely that *tudor5-prime* accomplishes this by tipping the balance of sexual identity in the germline towards a male program. This balance seems to be maintained by the additive effects of several genes whose expression is controlled by *Sxl*. *Phf7* was the first of these gene that was recently described to promote male identity in the germline and is repressed by Sxl in the female germline (Yang et al, 2012, Shapiro-Kunane et al, 2015).

While the expression of the other two genes in the female germline sex determination pathway, *ovo* and *otu*, are regulated both intrinsically by X chromosome dose, and extrinsically by signals from the soma (Nagoshi et al, 1995; Andrews and Oliver, 2002), Sxl's germline expression is controlled solely by the germline's X chromosome constitution (Hashiyama et al, 2011). In addition, the way in which ovo and otu promote female sexual identity in the germline is not known. This work as well as previous work from our lab and others pushes the field forward by proposing a molecular mechanism employed by SXL to promote female identity in the germline. This mechanism includes the

SXL-mediated repression of genes that promote male identity in the germline. SXL-mediated repression can occur at different levels. Repression at the transcriptional level, as seen with *tudor5-prime* and *Phf7* (discussed in chapter 2), likely occurs indirectly by SXL's action through/on another gene. SXL can also regulate genes at the translational level by binding directly to the target gene's mRNA. Preliminary evidence for SXL-mediated translational control of *tdrd5p* is reported here. Further experiments are needed to confirm this interaction.

The age-dependent appearance of the morphological defects in *tudor5-prime* mutant males, as well as their enhancement by high temperatures may suggest that *tdrd5p* is needed most in males under stressful conditions. It may also be that the function of tdrd5p can be compensated for by another gene; thus, the effect of its loss can be tolerated for some time. This suggests that tudor5-prime though potent enough to promote male identity in the germline, does not work alone. Additional experiments are needed to define the nature of the differentiation defect in *tdrd5p* mutant males. One important follow-up experiment is to analyze the expression of testis-specific transcription factors (tTAFs) such as *spermatocyte arrest (sa),* and meiosis I arrest (mia) in *tdrd5p* mutant testes. These tTAFs are turned on in spermatocytes and are required for meiotic cell cycle progression and spermatid differentiation (reviewed by White-Cooper, 2004). A defect in their expression would suggest failure of tdrd5p

mutant germ cells to properly enter meiosis, or a delayed entry into meiosis. This could explain the observed fecundity defect.

The molecular function of *tudor5-prime* is another topic ripe for study. I conducted several assays to determine whether *tdrd5p* is involved in the piRNA pathway, but did not find any evidence supporting this hypothesis. Another hypothesis is that *tudor5-prime* is involved in the mRNA decay or translational repression machinery. The localization of HA-tagged TDRD5P into cytoplasmic punctae is characteristic of ribonucleoprotein complexes (RNPs) involved in mRNA decay and translational repression (reviewed in Voronina et al, 2011). This suggests that TDRD5P may function with such an RNP. To test this hypothesis, I am using immunohistochemistry to assay for co-localization of TDRD5P:HA with well-known components of germline RNPs. To date, no significant co-localization has been found.

Current work on this project are studies to test SXL's direct regulation of *tudor5-prime*. This is being done using the same *tdrd5p*:HA BAC construct used in figures 3.2 and 3.3. The SXL binding site has been deleted from the intron and mutated in the 3'UTR of the *tudor5-prime* genomic DNA in this BAC. Two different mutants were made so that either the intronic binding site was missing from the construct or both sites were missing from one construct. The effect on TDRD5P expression using these mutants can be viewed by HA immunostain. It

would be very interesting to find that Sxl represses expression of *tudor5-prime,* both at the transcriptional and at the translational level.

**Table 3.1: Description of all *tudor5-prime* mutant alleles generated.**

The underlined portion of the wild type sequence is the guide sequence used. The bolded nucleotides represents the Protospacer Adjacent Motif (PAM), which is required for Cas9-meidated cleavage. Dashed lines represent deleted nucleotides. Red letters represent inserted nucleotides.

| Category 1: Mutants without a frame shift | | |
|---|---|---|
| **Wild type** | **AAGAAGCCGAGGATAACAATGATGTGCCGCTCGAT** | |
| Mutant Allele Name | Allele Sequence | Lesion Description |
| tdrd5p$^{1.11}$ | AAGAAG**CCG**A---------TGATGTGCCGCTCGAT | -9nt |
| tdrd5p$^{G21}$ | AAGAAG**CCG**------ACAATGATGTGCCGCTCGAT | -6nt |
| tdrd5p$^{J2}$ | AAGAAG**CCG**A---TAACAATGATGTGCCGCTCGAT | -3nt |
| tdrd5p$^{G7}$ | AAGAAG**CCG**A------ **ATGCCTCGGAAGAAGCCACAT**CAATGATGTGCCGCTCGAT | -6nt & + 21nt |
| tdrd5p$^{q4}$ | AAGAAG**CCG**A----AA**T**CAATGATGTGCCGCTCGAT | -4nt & + 1nt |
| tdrd5p$^{6.12}$ | AAGAAG**CCG**A------CAATGATGTGCCGCTCGAT | -6nt |
| tdrd5p$^{3.11}$ | AAGAAG**CCG**A----------TGTGCCGCTCGAT | -12nt |
| tdrd5p$^{q34}$ | AAGAAG**CCG**AGG---------ATGTGCCGCTCGAT | -9nt |
| **Category 2: Mutants with frame shift change to frame #2** | | |
| **Wild type** | **AAGAAGCCGAGGATAACAATGATGTGCCGCTCGAT** | |
| Mutant Allele Name | Allele Sequence | Lesion Description |
| tdrd5p$^{G16}$ | AAGAAG**CCG**A-----**ATGATGT**ACAATGATGTGCCGCTCGAT | -5nt & + 7nt |
| tdrd5p$^{1.14}$ | AAGAAG**CCG**A----- **ATGAAGAAGC**ACAATGATGTGCCGCTCGAT | -5nt + 10nt |
| tdrd5p$^{p2}$ | AAGAAG**C**----------AATGATGTGCCGCTCGAT | -10nt |
| tdrd5p$^{q26}$ | AAGAAG**CCG**AG-ATAACAATGATGTGCCGCTCGAT | -1nt |
| tdrd5p$^{F1}$ | AAGAAG**CCG**A--AT-----TGATGTGCCGCTCGAT | -7nt |
| tdrd5p$^{q5}$ | ----------GATGTGCCGCTCGAT | -34nt |
| tdrd5p$^{q28}$ | AAGAAG**CCG**AG----**C**CAATGATGTGCCGCTCGAT | -5nt & + 1nt |
| tdrd5p$^{J3}$ | AAGAAG**CC**----------AT | -25nt |
| tdrd5p$^{m3}$ | AAGAAG**CCG**AGG-TAACAATGATGTGCCGCTCGAT | -1nt |
| tdrd5p$^{m4}$ | AAGAAG**CCG**A-------GGTGATGTGCCGCTCGAT | -7nt |
| tdrd5p$^{3.7}$ | AAGAAG**CCG**---- **GCTTAATGATGGGCT**TAACAATGATGTGCCGCTCGAT | -4nt & + 15nt |
| tdrd5p$^{G1}$ | --AGAAG**C**--------ACAATGATGTGCCGCTCGAT | -10nt |

| Category 3: Mutants with frame shift change to frame #3 | | |
|---|---|---|
| **Wild type** | **AAGAAGCCGAGGATAACAATGATGTGCCGCTCGAT** | |
| Mutant Allele Name | Allele Sequence | Lesion Description |
| tdrd5p$^{m7}$ | AAGAAG**CCG**----------CTCGAT | -20nt |
| tdrd5p$^{G11}$ | AAGAAG**CCG**A-----ACAATGATGTGCCGCTCGAT | -5nt |
| tdrd5p$^{G4}$ | AAGAAG**CCG**A--------ATGATGTGCCGCTCGAT | -8nt |
| tdrd5p$^{3.6}$ | AAGAAG**CCG**A-----<br>**TGTGATCATTACCCAATGATCAGCGGC**ACAATGATGTGCCGCTCGAT | -5nt & + 27nt |
| tdrd5p$^{E6}$ | AAGAAG**CCG**A-----**AATT**TGATGTGCCGCTCGAT | -5nt & + 4nt |
| tdrd5p$^{m6}$ | AAGAAG**CCG**A----**AC**AACAATGATGTGCCGCTCGAT | -4nt & +2nt |
| tdrd5p$^{4.15}$ | AAGAAG**CCG**AGG-**CG**TAACAATGATGTGCCGCTCGAT | -1nt & +2nt |
| tdrd5p$^{m9}$ | AAGAAG**CCG**A---**TGTCATA**TAACAATGATGTGCCGCTCGAT | -3nt & + 7nt |

# Chapter 4:

# Screen for sex-specific gene isoforms with roles in the undifferentiated germline

## Introduction:

Alternative Exon Usage

Alternative exon usage is a molecular mechanism used to create different isoforms of a gene. This is accomplished by using different combinations of the exons in a gene to create unique transcripts. Exon usage is influenced by promoter choice, which creates multiple pre-mRNAs from the same gene, as well as by alternative splicing, which creates different mature mRNAs from a single pre-mRNA transcript. These mechanisms are frequently used in eukaryotes to increase the diversity of proteins. In fact, 92-94% of human genes are alternatively spliced (Wang et al, 2008). However, their numerous roles in regulating developmental biology processes underscores their importance past mere proteome diversity. Examples of this can be found in the *Drosophila* sex determination pathway. A classic example is the creation of a male-specific and a female-specific isoform of the DSX protein (described in Chapter 1). These two proteins have distinct and opposing functions—the female isoform promotes female sexual identity in the soma, while the male isoform promotes male identity.

## Alternative Exon Usage in Undifferentiated Gonads

In *Drosophila*, several studies have established exon usage as a mechanism that is frequently used to produce tissue-specific as well as sex-specific gene isoforms (McIntyre et al., 2006; Telonis-Scott, 2009). Interestingly, sex-specific gene isoforms have also been found in gonads enriched for the undifferentiated germline (Gan et al., 2010). This suggests that alternative exon usage may play a role regulating germline sexual identity. Q. Gan's data showing sex-specific gene isoforms in the undifferentiated germline was generated by an RNA-Seq experiment done in the lab of Dr. Xin Chen. This RNA-Seq experiment was done using male and female adult gonads from *bag of marbles* (*bam)* mutant flies. The *bam* gene is required for germline differentiation—germ cells in *bam* mutant animals are unable to transition from the proliferation to the differentiation stage ( Gönczy, Matunis, and DiNardo, 1997, McKearin and Spradling, 1990). As a result, the use of *bam* mutants allows for the enrichment of genes needed in the early (undifferentiated) germline. Their analysis of the *bam* mutant male and female dataset showed that a subset of genes produce sex-specific isoforms. In addition, Gan et al found an enrichment of splicing factors in these undifferentiated gonads. Alternative splicing is the primary mechanism used for the establishment of sexual identity in the soma, and this data provided evidence

in support of alternative splicing as an important mechanism in germline development as well.

Female germline sexual identity is dependent on the action of SXL in the germline. *Sxl* is required in the female germline not only for sexual identity but also for proper germ cell development. SXL is an RNA binding protein that can regulate its own expression as well as the expression of other genes, at the level of splicing or at the translational level. It regulates gene expression at the translational level specifically by binding to the 3'UTR to repress translation of that gene. This mechanism of action of SXL has been shown with *msl2* in dosage compensation (Kelley et al, 1997), and more recently shown for regulating *nanos* function in the female germline (Chau et al, 2012). SXL's involvement in pre-mRNA splicing is key in somatic sex determination. Owing to this, SXL's specific role has been very well studied in somatic cells but largely ignored in germ cells. The pronounced use of alternative gene isoforms to decide between the male and female developmental programs in the soma led us to hypothesize that the use of sex-specific gene isoforms—created either through SXL's alternative splicing function, or by the action of some yet unidentified gene—is a mechanism that is also used to establish and maintain sexual identity in the germline. In this chapter I discuss my investigation into the role that alternative gene isoforms play in determining and maintaining germline sexual identity as well as in

regulating proper germline function. I discovered that while many exons may be sex-specifically expressed in the germline as can be seen by RNA-Seq and verified by RT-PCR, that is not necessarily a biologically significant event for sex-specific germ cell development.

## Materials and Methods

### Fly Stocks

The fly stocks used were obtained from Bloomington Stock Center unless otherwise indicated. The TRiP RNAi lines used are listed in table 4.3. nanos-GAL4-VP16 (Van Doren, 1998). All RNAi crosses were done at 29$^{o}$C.

### Immunofluorescence

Adult ovaries and testes were fixed, blocked and stained as described in Gonczy et al, 1997. All images were taken with a Zeiss LSM 510 Confocal microscope. Primary antibodies and the concentrations used are as follows: chicken anti-Vasa 1:10000 (K. Howard); rabbit anti-vasa 1:10,000 (R. Lehmann); rat anti-Ncadherin 1:12 (DN-EX#8, DSHB); mouse anti-HTS 1:4 (1B1, DSHB); mouse anti-armadillo 1:100 (N2 7A1, DSHB) DSHB). DSHB: Developmental Studies Hybridoma Bank. Secondary antibodies were used at 1:500 (Alexa-fluor).

## RT-PCR

Total RNA was isolated from *bam* mutant ovaries and testes using RNA-bee (Tel-Test). Contaminating DNA was removed from the RNA using Turbo-DNA-free (Ambion). RNA was converted to cDNA using Superscript II (Invitrogen). PCR amplification was done using gene-specific primers listed (Appendix C). Primers were designed to cross exon boundaries of candidate sex-specific exons, and primers across exons with similar expression between the 2 sexes were used as controls. Annealing temperature for PCRs ranged from $56^0$C – $60^0$C. Sex-specific expression was scored based on on/off expression as well as reduction in levels compared to the control. PCR validation for each exon was done twice using two biological replicates.

## Creating transgenic RNAi lines

Short hairpin RNAi constructs were designed and cloned according to the protocol described by the Transgenic RNAi Project (TRiP), (Ni et al, 2011) (see table 4.3 for RNAi sequences). RNAi constructs were designed against sex-specific exons as well as common exons as controls and cloned into the VALIUM 22 or VALIUM 20 vectors. Constructs were injected by Genetic Services Inc. All constructs were integrated into the Drosophila genome using phiC31-mediated site-specific integration into the attP2 landing site. Screening of the resulting

adult flies for transformants was done using the $y^1v^1$/attP2/TM3 Sb stock to establish stable transgenic lines.

## Results

### Candidate genes display complex patterns of sexually dimorphic exon expression

This work was initiated in 2010, before I conducted the RNA-Seq experiment described in chapter 2. Therefore, the RNA-Seq data used for this project, comparing the expression profiles of bam mutant males and bam mutant females, was kindly provided by Qiang Gan and Xin Chen (Gan et al, 2010). I applied filtering parameters to the RNA-Seq data to reduce noise in the data and enrich for those exons with highest sex-specific expression.

Four key filtering parameters were applied:

1. Gene expression is greater than background (RPKM >=10)
2. Exon expression is greater than background (RPKM >=10) in the relevant sex
3. Exon usage is highly sex-specific (4-fold or more difference in expression)

4. Exon expression difference is the result of sex-specific exon usage not sex-specific gene expression (i.e exon expression difference is 2 fold above gene expression difference between male and female).

Expression levels were assessed based on RPKM (total reads per kilobase of gene model per million Mapped reads) of each exon or gene. This data filtering step yielded 358 exons covering 242 genes as having 4 fold or higher sexually dimorphic expression (Figure 4.1). One potential limitation of applying the above filtering parameters is biasing against sex specifically spliced exons in genes with already sexually dimorphic expression in that same sex. However, one major advantage of the filters is that they allow us to enrich for sexually dimorphic genes, with alternative splicing events that cause sex-specific expression of exons in the opposite sex. *Sxl* is a prime example of this. It is expressed more highly in females, however the exon that is spliced out in females and retained in males is captured in our screen as showing higher expression in males.

To further eliminate any false positives that escaped the filters, I manually evaluated the expression pattern of each candidate exon using the UCSC genome browser display of the RNA-Seq data. Then I evaluated each exon usage event and categorized it as either alternative splicing or alternative promoter usage (Table 4.1). Finally, I prioritized the order of the screen using subjective and objective criteria and gave each exon a rank from 1 to 10. The Criteria used were:

gene function, whether or not coding region is changed by the alternative splicing event, complexity of the splicing pattern of the gene, exon shows clear dimorphic expression, and all genes of unknown function (20%) were ranked highly. I chose to validate the exons with the highest ranking, of 8 and above by RT-PCR. This resulted in a list of 148 exons covering 83 genes.

Many of these 88 genes exhibited very complex and interesting splicing patterns. More than half of the alternative exon usage events in these genes resulted directly from alternative promoter usage events. This exemplifies the importance of alternative promoter usage in transcript diversity and it likely also plays a large role in the expression of sex-specific gene isoforms. In fact this is known to happen in males, as there are testis specific transcription factors, tTAFs, (TBP-associated factors) that are vital for male-specific transcription in the testis. The idea of sex-specific promoter usage is reflected well in this dataset. It will be interesting to find other genes that like Sxl, use different promoters to regulate sex specific gene expression. Of the 148 exons, 57 of them showed higher expression in males, while 91 of these exons, over 60%, were expressed more highly in females. The set of genes with female-specific exons contained genes such as *bunched*, *paxilin*, and *14-3-3ζ*, already known to have important functions in gonad function, though a role for

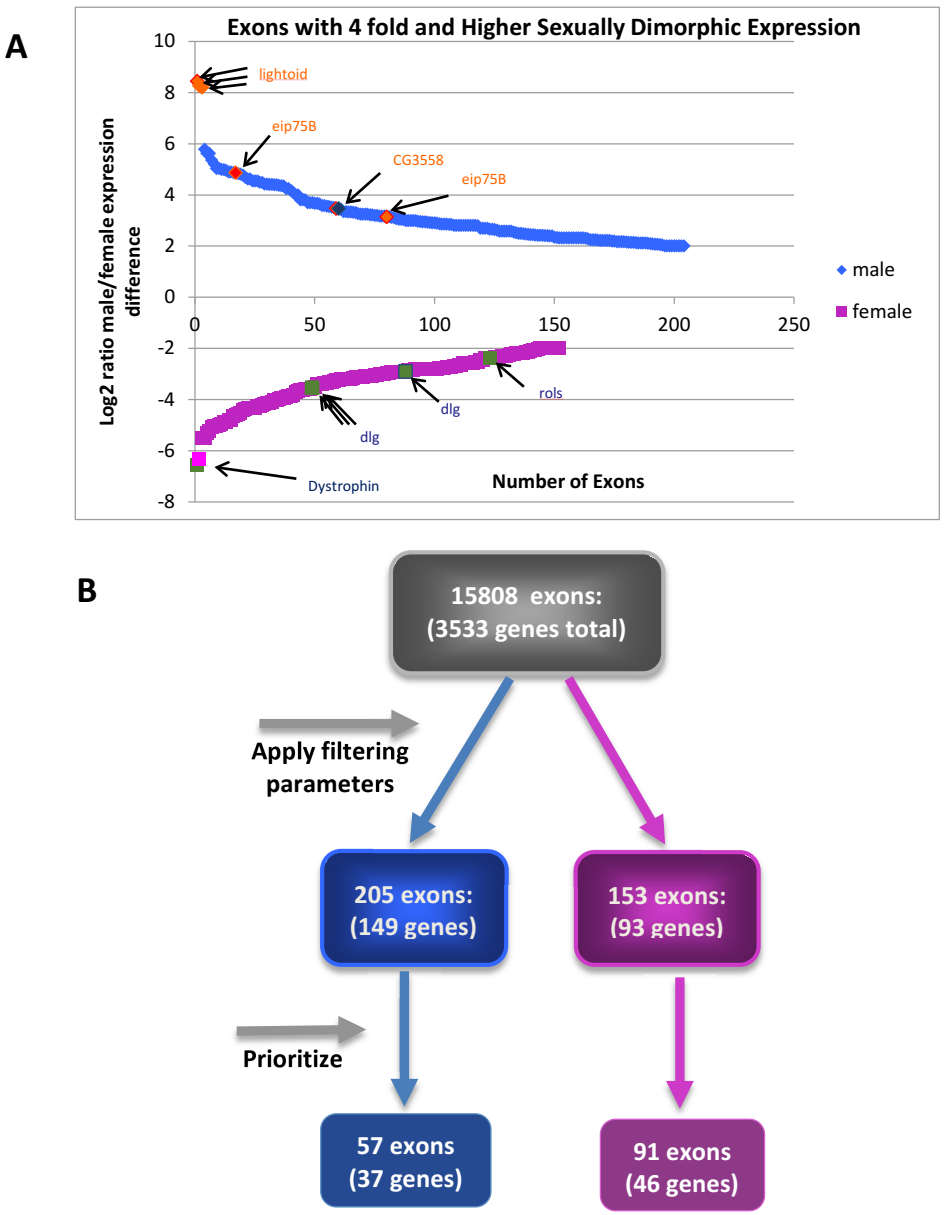



**Fig 4.1: RNA-Seq shows Sex-specifically expressed Exons**. RNA-Seq data kindly provided by Qiang Gan and Xin Chen (Gan et al, 2010). A). Exons with 4 fold and higher dimorphic expression. B) Flow chart showing change in number of exons by filtering of the RNAseq data and prioritizing to obtain a list of 155 exons

alternative splicing had not been explored. A few genes, for eg *LpR2, jbug* and *Zasp* (aka *Zasp52)* had exons that were dimorphically expressed in both sexes. The *Zasp* gene in particular has a very complex transcript structure, having 18 different transcripts transcribed from four different promoters. Additionally, the RNA-Seq data for *Zasp* showed four exons expressed more highly in males and five expressed more highly in females. This complexity in transcript structure may be important for proper regulation of this gene. And the sex-specific exon expression patterns could be a result of this regulation.

## Alternative exon usage is confirmed as sex specific in half of the events tested

I validated the RNA-Seq expression data of the 148 selected alternative exon usage events by doing RT-PCR on *bam* mutant male and female gonads. Primers were designed to amplify across exon boundaries of exons to be tested. A set of exons displaying similar levels of expression (according to RNA-Seq) in both sexes was chosen for each gene. These were used as controls for the normal sex expression-difference of that gene. One disadvantage of using RT-PCR for this step is that it is not quantitative, therefore, there is a risk of missing exons that are not completely unused in a particular sex, but rather is simply not used as much. In addition, since splicing is not 100% efficient, there may be some alternative splicing events that produce a few transcripts of the sex-specific

isoform in the opposite sex. This would be a problem especially if that particular gene is normally expressed at a low level in both sexes. In this case, the exponential amplification cycles of non-quantitative PCR could result in product expression that at plateau appears similar in both sexes, resulting in a loss of that original sexually dimorphic exon expression. To account for this possibility, PCRs were done using a low cycle number (26 cycles). In addition, sex-specific expression was not scored based solely on on/off expression but also based on reduction in levels compared to the control.

I completed the RT-PCR validation with the striking result that half of the exons show sex-specific expression by RT-PCR (Table 4.1). Specifically, 48 (53%) of the 91 exons that showed female-specific expression by RNA-Seq, also showed a female-specific expression pattern by RT-PCR. And 32 of 57 exons, 56%, resulted in a male-specific expression pattern. 15% of the PCRs did not work. Interestingly, six genes were validated as having both male and female sex-specific exons. The *LpR2* gene falls into this category; all of its predicted male-specific as well as the female-specific exons were verified. Furthermore, the use of sex-specific promoters in many genes was confirmed.

**Table 4.1: Results of RT-PCR validation of sex-specific exon usage**. APU: alternative promoter usage. AS: alternative splicing. n/a: not available (PCR did not work). Blue shaded exons are more highly expressed in males and non-shaded are higher in females.

| Gene Symbol | Exon Name | Usage Category | PCR result (sex-specific: yes/no) |
|---|---|---|---|
| 14-3-3zeta | CG17870.13 | AS | no |
| Asph | CG8421.8 | AS | no |
| Atpalpha | CG5670.17 | AS | no |
| betaTub97EF | CG4869.5 | AS | no |
| bun | FBgn0010460.5 | APU | yes |
| ced-6 | CG11804.9 | APU | no |
| CG11961 | CG11961.1 | APU | yes |
| CG12065 | CG12065.5 | AS | yes |
| CG17370 | CG17370.9 | APU | n/a |
| CG31522 | CG31522.271968-271888 | AS | n/a |
| CG32560 called CG42684 in flybase and CG42270 in UCSC | FBgn0052560.3 | AS | no |
| CG33275 | FBgn0035802.4 | APU | no |
| CG33995 - same as CG31919 exon 13 | CG33995.2 | AS | n/a |
| CG33995 - same as CG31919 exon 19 | CG33995.14 | APU | yes |
| CG34394 | FBgn0085423.3 | AS | n/a |
| CG34417 | FBgn0085446.13 | AS | yes |
| CG34417 | FBgn0085446.8 | AS | yes |
| CG34417 | FBgn0085446.9 | AS | yes |
| CG3558 | CG3558.6 | AS | yes |
| CG5850 | FBgn0032172.4 | AS | no |
| CG5973 | CG5973.1 | APU | no |
| CG5973 | CG5973.9 | APU | yes |
| CG6043 | FBgn0032497.1 | APU | no |
| CG6043 | FBgn0032497.15 | AS | no |
| CG6043 | FBgn0032497.16 | AS | no |
| CG6043 | FBgn0032497.2 | APU | no |
| CG6043 | FBgn0032497.3 | AS | no |
| CG6145 | CG6145.5 | APU | yes |
| CG6767 | CG6767.9 | AS | yes |
| CG7378 | FBgn0030976.3 | AS | no |
| CG7852 | CG7852.1 | APU | yes |
| dlg1 | FBgn0001624.10 | APU | yes |
| dlg1 | FBgn0001624.11 | APU | yes |
| dlg1 | FBgn0001624.2 | APU | yes |
| dlg1 | FBgn0001624.3 | APU | yes |
| dlg1 | FBgn0001624.4 | APU | yes |
| dlg1 | FBgn0001624.6 | APU | yes |
| dlg1 | FBgn0001624.7 | APU | yes |

| | | | |
|---|---|---|---|
| dre4 | CG1828.16 | AS | no |
| Dys/ det | FBgn0024242.20 | AS | yes |
| Ect4 | FBgn0085402.10 | APU | yes |
| Ect4 | FBgn0085402.18 | APU | yes |
| Ect4 | FBgn0085402.9 | APU | yes |
| eIF4G | FBgn0023213.10 | AS | n/a |
| eIF4G | FBgn0023213.17 | APU | n/a |
| Eip75B | CG8127.2 | APU | yes |
| Eip75B | CG8127.3 | AS | yes |
| Eip75B | CG8127.7 | APU | yes |
| ens/CG14998 | FBgn0035500.9 | AS | no |
| Fas3 | FBgn0000636.8 | AS | no |
| Fim | CG8649.8 | APU | no |
| fus | CG8205.7 | AS | yes |
| fz2 | FBgn0016797.3 | APU | n/a |
| gce/CG15032/ CG42739 | FBgn0030626.1 | APU | n/a |
| hts | CG9325.12 | AS | n/a |
| Ih | CG8585.16 | AS | yes |
| Ih | CG8585.2 | APU | yes |
| jbug | CG30092.17 | APU | yes |
| jbug | CG30092.19 | APU | yes |
| jbug | CG30092.3 | APU | yes |
| jbug | CG30092.4 | APU | yes |
| jbug | CG30092.5 | APU | yes |
| jbug | CG30092.7 | APU | yes |
| jbug | CG30092.9 | APU | yes |
| l(1)G0232 | CG32697.8 | APU | n/a |
| loco | CG5248.2 | APU | yes |
| loco | CG5248.6 | APU | yes |
| LpR2 | FBgn0051092.11 | APU | yes |
| LpR2 | FBgn0051092.12 | APU | yes |
| LpR2 | FBgn0051092.13 | APU | yes |
| LpR2 | FBgn0051092.14 | APU | yes |
| LpR2 | FBgn0051092.4 | AS | yes |
| LpR2 | FBgn0051092.9 | AS | yes |
| ltd | CG8024.1 (my Add) | APU | yes |
| ltd | CG8024.10 | APU | yes |
| ltd | CG8024.5 (my Add) | APU | yes |
| ltd | CG8024.6 (my Add) | APU | yes |
| ltd | CG8024.7 | APU | yes |
| ltd | CG8024.8 | APU | yes |
| Mctp | FBgn0034389.1 | APU | yes |
| Mctp | FBgn0034389.5 | AS | yes |
| Mical | CG33208.10 | AS | yes |

| | | | |
|---|---|---|---|
| Mical | CG33208.24 | AS | yes |
| mth | CG6936.1 | APU | n/a |
| mth | CG6936.8 | AS | n/a |
| Nhe2 | CG9256.13 | AS | no |
| Pax | CG31794.2 | AS | yes |
| Pax | CG31794.35 | APU | yes |
| Pax | CG31794.37 | AS | yes |
| PHGPx | CG12013.6 | AS | no |
| Pino | CG4710.2 | APU | no |
| pnt | CG17077.11 (MVD Add) | APU | n/a |
| pnt | CG17077.12 (MVD Add) | APU | n/a |
| pnt | CG17077.3 (MVD Add) | AS | yes |
| pnt | CG17077.4 (MVD Add) | AS | yes |
| ps | FBgn0026188.9 | AS | no |
| Pvr | CG8222.4 | AS | yes |
| Rbp9 | CG3151.3 | APU | no |
| rols | CG32096.6 | AS | yes |
| scb | CG8095.9 | APU | yes |
| shep | CG32423.3 | AS | no |
| shot | CG18076.2 | AS | yes |
| shot | CG18076.23 | AS | yes |
| shot | CG18076.27 | AS | yes |
| shot | CG18076.28 | AS | yes |
| shot | CG18076.30 | AS | yes |
| shot | CG18076.35 | APU | no |
| sun | CG9032.4 | AS | no |
| Tm1 | CG4898.17 | AS | no |
| Tm1 | CG4898.3 | AS | no |
| Tm1 | CG4898.5 | AS | no |
| Tm1 | CG4898.6 | AS | no |
| Tm1 | CG4898.8 | AS | no |
| tmod | FBgn0082582.1 | APU | yes |
| tmod | FBgn0082582.3 | APU | n/a |
| tmod | FBgn0082582.4 | APU | yes |
| tmod | FBgn0082582.6 | AS | yes |
| Treh | CG9364.10 | APU | no |
| Treh | CG9364.11 | APU | no |
| Treh | CG9364.12 | AS | no |
| Treh | CG9364.13 | APU | no |
| trio | CG18214.22 | APU | yes |
| trio | CG18214.23 | APU | yes |
| trol | FBgn0001402.1 | APU | n/a |

| | | | |
|---|---|---|---|
| Trp1 | CG4758.5 | APU | no |
| ttk | CG1856.5 | AS | no |
| ttk | CG1856.7 | APU | yes |
| ttk | CG1856.8 | APU | yes |
| tws | CG6235.2 | APU | no |
| Vha44 | CG8048.10 | APU | n/a |
| VhaSFD | CG17332.7 | AS | n/a |
| vir-1 | FBgn0043841.1 | AS | n/a |
| vir-1 | FBgn0043841.2 | AS | n/a |
| wls | CG6210.2 | AS | yes |
| yki | FBgn0034970.5 | AS | no |
| Zasp aka zasp52 | FBgn0083919.11 | AS | yes |
| Zasp aka zasp52 | FBgn0083919.12 | AS | yes |
| Zasp aka zasp52 | FBgn0083919.13 | AS | yes |
| Zasp aka zasp52 | FBgn0083919.14 | AS | yes |
| Zasp aka zasp52 | FBgn0083919.15 | APU | no |
| Zasp aka zasp52 | FBgn0083919.25 | APU | no |
| Zasp aka zasp52 | FBgn0083919.26 | APU | no |
| Zasp aka zasp52 | FBgn0083919.6 | AS | no |
| Zasp aka zasp52 | FBgn0083919.8 | AS | no |
| zfh1 | CG1322.4 | APU | no |
| zfh1 | CG1322.5 | APU | n/a |
| zfh1 | CG1322.6 | APU | n/a |
| zip | CG15792.7 | AS | yes |

## Knockdown of the sex-specific isoforms produced mostly mild phenotypes

The final stage of the project was to determine whether or not the sex-specific exons belonged to transcripts that had biologically significant functions in the germline of the respective sex. To accomplish this, I did a loss of function screen using RNAi to knock down sex-specific isoforms specifically in the germline. In addition to knocking down sex-specific transcripts, common transcripts were knocked down by targeting common exons, using either RNAi constructs that I made or commercially available RNAi lines. To achieve transcript-specific knockdown I used the VALIUM vector system, which uses short hairpin microRNA-like (shmiRNA) constructs for RNA interference (Ni et al, 2011). Since these short hairpin RNAi constructs are only 21nt long, they can be used to target specific exons to knock down specific isoforms of a gene. This method provides the unique capability to assess the specific contribution of select gene isoforms in a way that null alleles and deficiencies cannot be used.

I created a total of 29 transgenic shmiRNA fly lines covering 23 genes. All RNAi lines were expressed in the germline using a *nanos*-Gal4 driver. To assay for phenotypes I used immunohistochemistry on adult gonads, looking for defects in germ cell development and differentiation. Unfortunately these lines did not yield robust phenotypes (Fig 4.3, table 4.2). Germ cell loss was a frequent phenotype that was observed especially in males, but this phenotype was usually

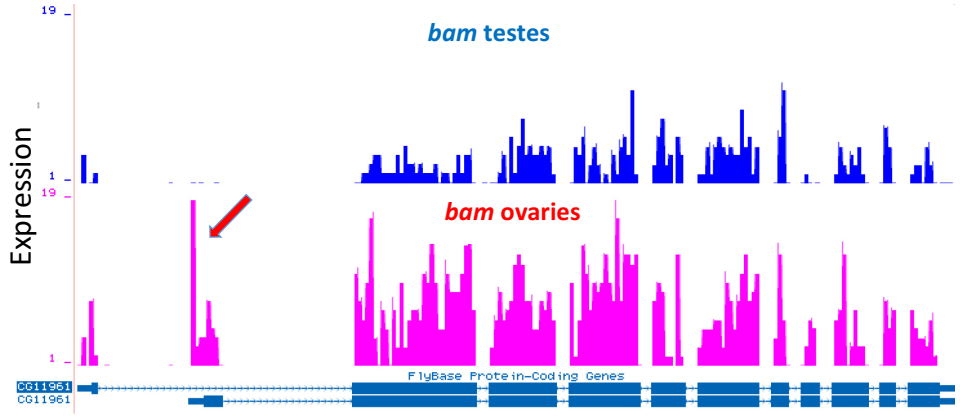
mild or low penetrance. One of the only RNAi knockdowns which gave a female phenotype was for a gene named *zipper*. *zipper* makes a non-muscle myosin heavy chain protein (Bloor and Kiehart, 2001). It's been shown to be important for border follicle cell migration in the ovary (Liu & Montell, 1999; Dobens & Raftery, 2000). RNAi against the supposed male-specific exon (exon 7) of this gene caused germline loss in both males and females (Fig 4.3 G-H), however, which suggests that these transcripts are also used in females. Since the germline loss phenotype is not as severe in females, it is possible that the set of transcripts containing exon 7 are used less in females compared to males. This explains the male-specific RNA-Seq result for exon 7. The knockdown of all *zipper* isoforms using a different RNAi line produced severe germline loss in both sexes (Fig 4.3 I-J), confirming the importance of this myosin heavy chain gene in the germline of both males and females.

RNAi phenotypes being produced in the opposite sex than what was expected based on the RNA-Seq data was a common observation. One example of this is the gene *CG12065*, a nucleoside phosphorylase involved in lateral inhibition (Mummery-Widmer et al, 2009). Mild male germline loss occurred in RNAi against the sex-specific exon as well as a common exon of this gene (Fig4.3D). Since the RNA-Seq data showed that sex-specific exon 5 was more highly expressed in females, this is likely a situation similar to *zipper*; the

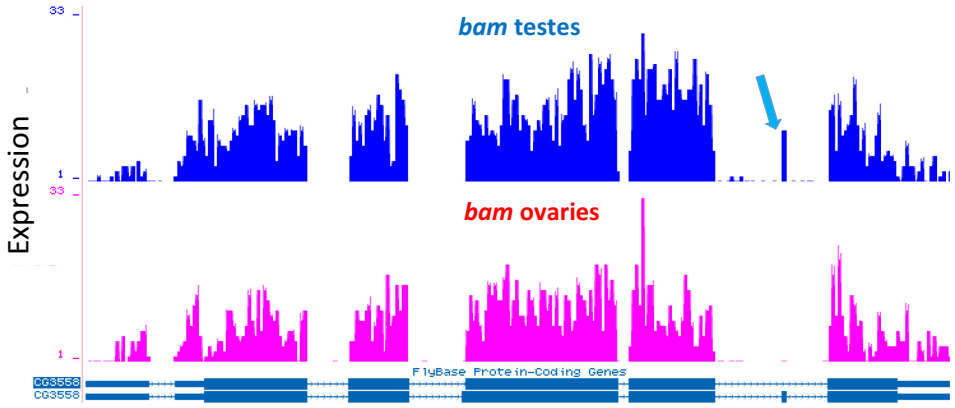
**Fig 4.2: Sample genes with sex specific exons**. A-B) UCSC genome browser display of bam male and female RNA-Seq data (Q. Gan) for CG11961 and CG3558 which have 1 female and 1 male-specific exon respectively. C) UCSC genome browser display showing jbug which has sexually dimorphic exon expression in both sexes.

A

Genes with sex-specific exon expression in only one sex

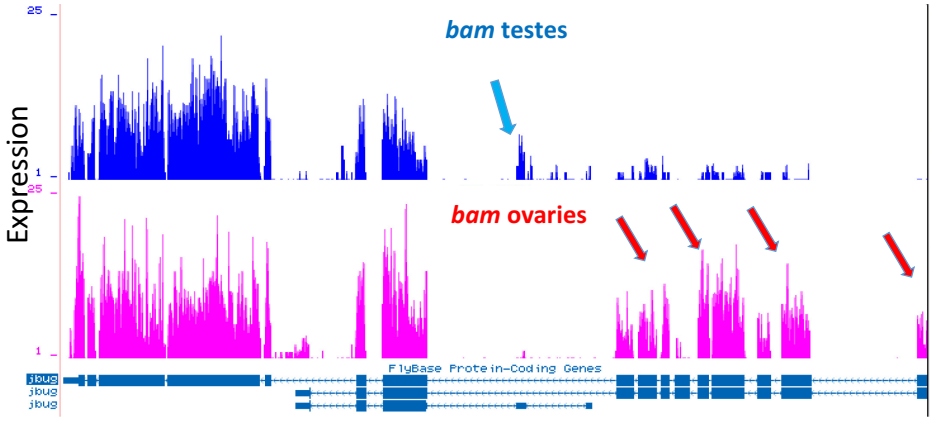



B




C

Gene with sex-specific exon expression in both sexes



transcripts containing exon 5 are used less in males but still required. *rols* (*rolling pebbles*) is an adaptor protein shown to be involved in border follicle cell migration (Borghese et al, 2006). Knockdown of this gene using RNAi against a common exon caused reduced sperm production (Fig 4.3L), suggesting that this gene is important for efficient germline differentiation into sperm. One phenotype that was mild though quantifiable was for the *bun* (*bunched*) gene. This gene is a transcription factor that is downstream of dpp signaling. Exon 5 was more highly expressed in female *bam* RNA-Seq, but knockdown of *bun* produced a phenotype only in males, a mild germline loss represented by shallow testes (Fig 4.3M).

In addition to designing shmiRNA constructs to create transgenic RNAi flies, I also used commercially available VALIUM shmiRNA fly lines as they became available (created by the TRiP project, Ni et al, 2011). These lines were used to knock down the common isoforms of candidate genes. When available, TRiP lines were used in place of designing shmiRNAi to knock down whole gene function before knocking down sex-specific isoforms. Some TRiP lines were also used to confirm the results produced by RNAi lines previously created and analyzed. In total I knocked down 20 genes by germline-specific RNAi using TRiP lines and loss of 8 of these genes resulted in gonad phenotypes (table 4.3). Phenotypes range from mild to severe, and similar to what I observed with the

RNAi lines that I designed, most of the phenotypes appeared in testes. The gene *loco* (*locomotion defects*) is one of the few genes that produced a striking phenotype; it caused severe germline loss in aged males (15 days old: Fig 4.3F). *loco* is a GTPase activator with male semi-sterile alleles.

One of my first RT-PCR validated candidates is a gene called *discs large* (*dlg1*). The RNA-Seq data for *dlg1* in *bam* mutants suggests the existence of two distinct sets of transcripts. One set beginning at exon 1, seems to be female-specific (figure 4.4F red arrows), and the second set beginning at exon 13, seems to be commonly expressed in both sexes. This gene appears to represent a classic example of the use of alternative promoter usage to create sex-specific gene isoforms. The existence of these two sets of transcripts as well as their sex-specific expression pattern was validated by RT-PCR (Figure. 4.4E). However, analysis of the protein expression of these transcripts using GFP-tagged protein trap lines revealed that the predicted female-specific transcripts are expressed in males as well as females. In fact, these transcripts are not expressed in the germline at all. Instead, they are expressed in the hub and the very early somatic cells in males, and throughout the soma in the follicle cells in females (Fig 4.4A-B). This expression pattern—in many cells in females and few cells in males—is what gave this set of transcripts a sex-specific expression pattern in the RNA-Seq data.

**Fig 4.3: Sample of RNAi phenotypes from the Screen.** A-F) Confocal images of wild type gonads and specified RNAi knockdown gonads..C-D) *CG12065* RNAi against either the sex-specific or the common exon. E-F) *loco* RNAi against the common exon. G-H) *zip* RNAi against a exon 7 I-J) zip RNAi against the common exon. K-L) Confocal imagse of Dapi marking condensed sperm heads (arrows) in control and *rols* RNAi testes. The *rols* RNAi testes have less sperm. M) Quantification of testis depth in control and *bun* RNAi testes. Shallow testes is an indication of germ cell loss. All RNAi done using nanos Gal4 driver at 29$^{o}$C. Driv only ctrl: nanos-GAL4 driver only Control.

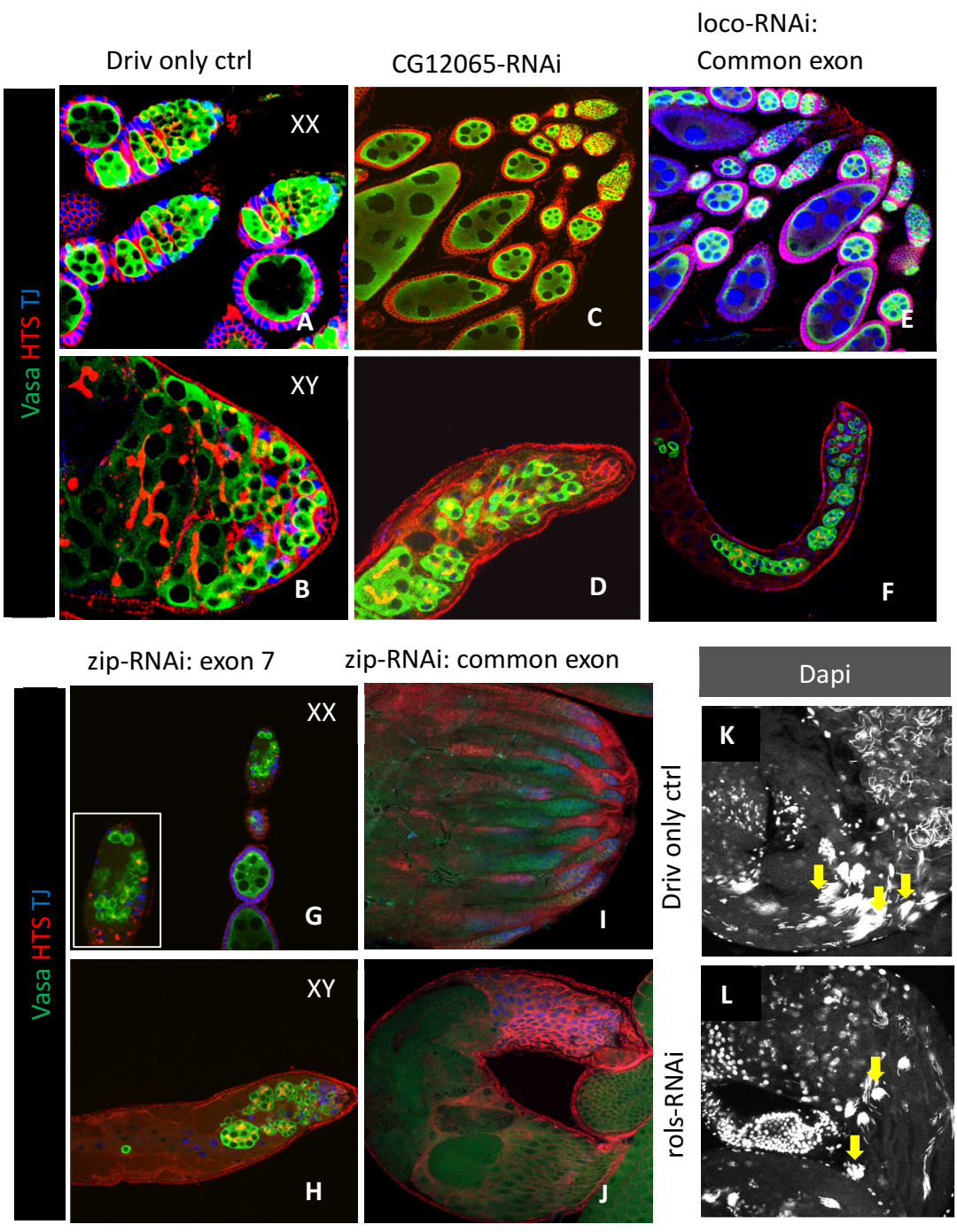

Driv only ctrl
CG12065-RNAi
loco-RNAi:
Common exon
Vasa HTS TJ
XX
XY
A
B
C
D
E
F
zip-RNAi: exon 7
zip-RNAi: common exon
Dapi
Vasa HTS TJ
XX
XY
G
H
I
J
K
L
Driv only ctrl
rols-RNAi
M
Testis depth in uM
20
15
10
5
0
Driv only ctrl
bun RNAi

**Table 4.2: Summary of shmiRNAi lines used to knock down sex-specific gene isoforms**. Lines were created following the protocol described in Ni et al, 2011. Shaded lines are designed against control (common) exons. The TRiP dlg1 shmiRNA constructs were cloned into the opposite vector to germline loss phenotype and determine the effect of each vector on RNAi efficacy in the germline.

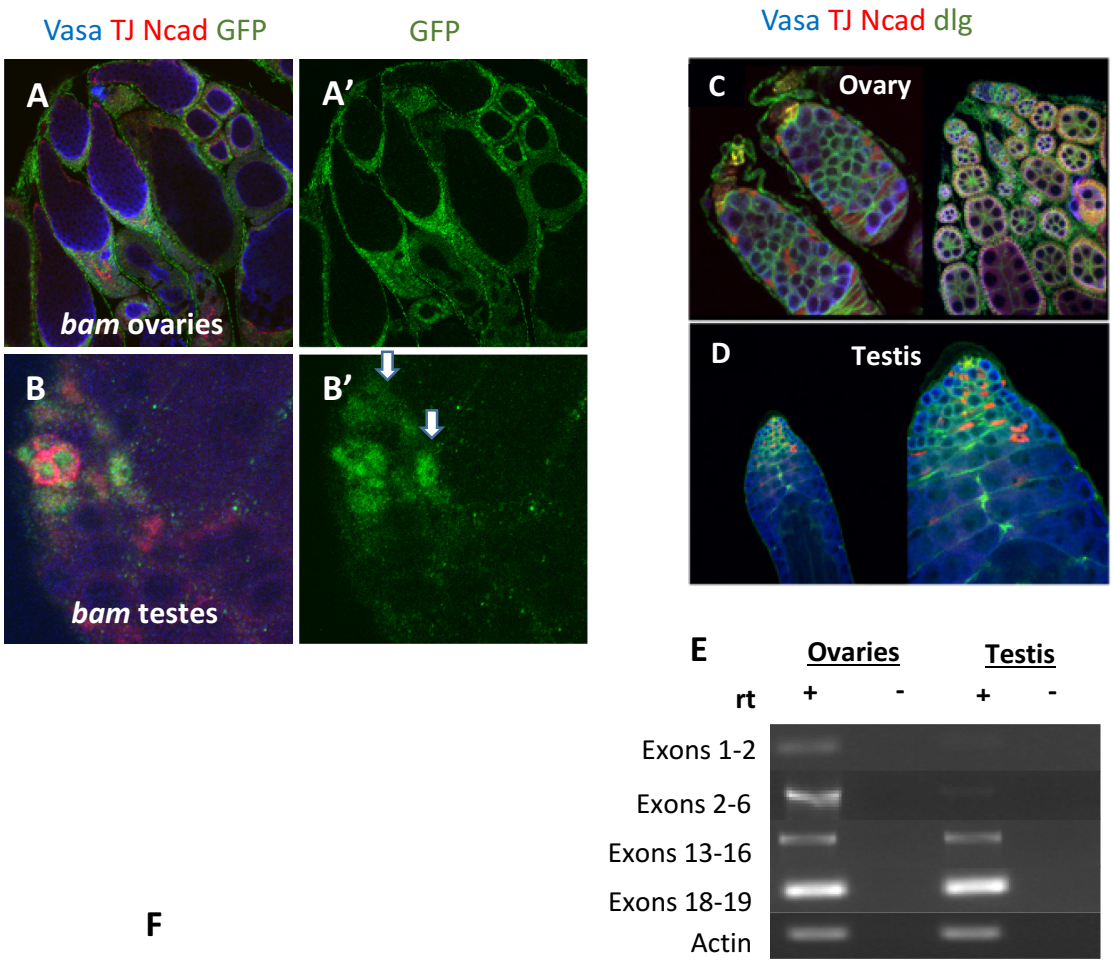
| Gene Symbol | Targeted Exon Name | Vector | 21bp Sequence | Phenotype |
|---|---|---|---|---|
| bun | FBgn0010460.1 | VALIUM 20 | ATCGATCAATGTTGA ATGTTT | Shallow testes. Male germline loss (mild) |
| CG11961 | CG11961.1 | VALIUM 22 | GACCAACGATCCTA CGAATGT | none |
| CG11961 | CG11961.6 | VALIUM 22 | AGTTGTTCAACAAA TCGAAGA | none |
| CG12065 | CG12065.5 | VALIUM 22 | GGTACATCATCTAC TGCTACG | Male germline loss |
| CG12065 | CG12065.7 | VALIUM 22 | AGGAGATTAAGGT GGTCAAGC | none |
| CG34417 | FBgn0085446.13 | VALIUM 20 | TAGCGATGCTAGCG AAGCGAA | none |
| CG3558 | CG3558.3 | VALIUM 20 | CCCGGAGCTGACGC TGTTCAT | none |
| CG5973 | CG5973.6 | VALIUM 20 | TTGCTGGATTACAT TCAAGAA | none |
| dlg1 | Fbgn0001624.2, Fbgn0001624.3, FBgn0001624.4 | VALIUM 22 | CTCGAAAGCGACTCC AGTAAA | none |
| dlg1 | FBgn0001624.18 | VALIUM 22 | CGGCACCGACACCTC CATCTA | Male germline loss. Egg chamber defect |
| dlg1 | FBgn0001624.18 | VALIUM 20 | TRiP GL00188 line shmir (orig in v22 vector) | Male germline loss |
| dlg1 | FBgn0001624.18 | VALIUM 22 | TRiP HMS00024 line shmir (orig in v20 vector) | Male germline loss |
| Dys / det | FBgn0024242.20 | VALIUM 20 | CCGCACTGTTCAGC GGACTTA | none |
| ect4 | FBgn0085402.11 | VALIUM 20 | ATCGCAGACGATGAC CTCCAA | Male germline loss |
| Eip75B | CG8127.3 | VALIUM 22 | CTGGATGGAGCGGCG GAAGAA | none |
| Eip75B | CG8127.7 | VALIUM 22 | CTCGAAGTAGATCGA ACGAAA | none |
| Ih | CG8585.5 | VALIUM 20 | CACACGATGGATTGC CTTCAA | none |
| LPR2 | FBgn0051092.10 | VALIUM 20 | TCAACTTGCAGCCAG GAACAA | none |
| mctp | FBgn0034389.9 | VALIUM 22 | ATGGTATACGCTGAA GGACAA | none |
| mical | CG33208.16 | VALIUM 20 | AACGACTGTATGAGA CTCTTA | none |
| pax | CG31794.8 | VALIUM 20 | CCGAGATGTAGTTCT CCATAA | none |

| | | | | |
|---|---|---|---|---|
| pvr | CG8222.11 | VALIUM 20 | CAGGAGATTGACCAC ATTCAA | none |
| rols | CG32096.7 | VALIUM 20 | CTGGATGTTACGGCA GGCAAT | Reduced sperm. Male germline loss (mild) |
| shot | CG18076.10 | VALIUM 20 | TTGAACGACTTGAGT TGTTCA | none |
| tmod | FBgn0082582.10 | VALIUM 20 | CAGCTGATCGAGCAC ATCAAT | none |
| trio | CG18214.16 | VALIUM 20 | ACCGGCTTGATGATC AAGTTA | Male germline loss (mild) |
| wls | CG6210.1 | VALIUM 20 | GCCGAGAGATGTTTC CAGTAA | Male germline loss (mild) |
| zasp52 | FBgn0083919.2 | VALIUM 20 | GCGCGCGTGATTCTT GCAGAA | Male germline loss (mild) |
| zip | CG15792.7 | VALIUM 20 | CTTGTTTGTGTTGACC GAGAA | Male germline loss. Low penetrance female germline loss |

| TRiP # | Gene Symbol | Vector | Phenotype |
|---|---|---|---|
| JF02954 | bun | VALIUM 10 | none |
| HMS01111 | CG12065 | VALIUM 20 | Male germline loss |
| HMS00014 | dlg1 | VALIUM 20 | Severe male germline loss |
| GL00188 | dlg1 | VALIUM 22 | Male germline loss (mild) |
| HMS01521 | dlg1 | VALIUM 20 | Male germline loss (mild) |
| HMS01954 | dlg1 | VALIUM 20 | none |
| HMS00024 | dlg1 | VALIUM 20 | Severe male germline loss |
| JF01118 | Dys | VALIUM 1 | none |
| HMS01530 | Eip75B | VALIUM 20 | none |
| HMC04208 | fus | VALIUM 20 | none |
| JF03253 | lh | VALIUM 10 | none |
| HMS01990 | jbug | VALIUM 20 | Male germline loss (LP)<br>Late stage egg chamber degeneration (LP) |
| HMS00455 | loco | VALIUM 20 | Germline loss in aged males |
| JF01627 | LpR2 | VALIUM 1 | Germline loss in aged gonads (both) |
| JF02836 | ltd | VALIUM 10 | none |
| HMS01870 | ltd | VALIUM 20 | None |
| JF01625 | Mical | VALIUM 1 | none |
| JF03111 | Pax | VALIUM 10 | none |
| HMS01873 | scb | VALIUM 20 | none |
| JF02971 | shot | VALIUM 10 | none |
| JF01094 | tmod | VALIUM 1 | none |
| .JF02815 | trio | VALIUM 10 | Male germline loss (mild) |
| HMS03008 | ttk | VALIUM 20 | none |
| JF01133 | Zasp52 | VALIUM 1 | Male germline loss (mild) |
| HMS01618 | zip | VALIUM 20 | Severe germline loss (both sexes) |
| GL00623 | zip | VALIUM 22 | Severe male germline loss |

**Table 4.3: Summary of RNAi lines used from the TRiP project.** TRIP designs RNAi against exons common to all/ most transcripts and therefore these lines knock down all/most isoforms of the respective gene. Aged gonads are 15D old. LP: low penetrance.

**Figure 4.4: Expression of different dlg1 isoforms**. A-B) Expression of DLG1 protein made by predicted female-specific transcripts in ovaries and testes. Expression is monitored by GFP protein trap located between exons 6 & 7. Arrows point to GFP positive somatic cells around the hub. C-D) Expression pattern of DLG1 protein made by the predicted non sex-specific transcripts in ovaries and testes. Epitope for this antibody is in exon 18. E) RT-PCR validation of the sex-specific exon usage in dlg1. F) UCSC genome browser display of bam male and female RNA-Seq data (Q. Gan) for dlg1.


Vasa TJ Ncad GFP
GFP
Vasa TJ Ncad dlg
A
A'
bam ovaries
B
B'
bam testes
C
Ovary
D
Testis
E
Ovaries
Testis
rt + - + -
Exons 1-2
Exons 2-6
Exons 13-16
Exons 18-19
Actin
F
Scale
chrX:
10 kb
11270000
11275000
11280000
11285000
11290000
11295000
11300000
bam testes
bam ovaries
Expression
Exons
FlyBase Protein-Coding Genes
dlg1

## Discussion

This project did not uncover any interesting genes with sex-specific isoforms important for sex-specific germline development. The initial filtering and prioritization I used to select the exons to study may have caused me to inadvertently eliminate good candidates. Even though this was done as a way to eliminate false positives, and select exons with the highest expression differences in an unbiased way, it may have resulted in the elimination of real positives as well. One way that the filtering step of this project could be improved is by using more than just the RNA-Seq expression fold change to select candidates. There are several ways in which selection could be done using parameters that are more meaningful for studying spliceforms. Firstly, the sex-specific splicing of *Sxl* causes an exon which has a premature stop codon to be spliced out (see chapter 1, pgs 11-12 for full discussion). We could use this splicing model to select genes that follow a similar pattern expressing sex-specific exons with a premature stop. Secondly, exons which encode for specific domains have a higher likelihood of producing proteins with different functions from the proteins made by common transcripts. This is the case with dsx splicing; two sex-specific proteins are produced each with a different dimerization domain. And lastly, screening for genes that have SXL binding sites is a good way to select candidates which are also possible Sxl targets. This final method would bias against genes that are sex-

specifically spliced in a SXL-independent manner. General splicing factors, *spen* and *snf* have been shown as important for sex-specific *Sxl* splicing; they may be responsible for creating other sex-specific isoforms.

If we assume that in an unbiased selection of genes expressed in the gonads, that at least 10% of these genes should have important functions in the gonad, then the lack of robust phenotypes from my selection of genes may be an artifact of my analysis of gene function. One reason could be the method used to knock down gene expression. RNA interference is an imperfect method of knocking down gene expression, as knockdowns can sometimes be incomplete. To account for this limitation, I supplemented analysis of RNAi lines that I created with RNAi lines available from TRiP. My analysis of the two sets of RNAi lines complemented each other. In addition the VALIUM vector system is the most efficient system to date for RNAi expression in the Drosophila germline. During the course of this project, however, it became evident that the degree of knockdown depended on the specific VALIUM vector being used as well as shmiRNA efficacy. The *dlg* RNAi lines are a good example of this; most of them cause male germline loss but at a range of severity from mild to severe. And the HMS01954 line does not cause a phenotype at all (table 4.3). In addition, cloning the TRiP HMS00024 shmiRNA construct into the VALIUM 22 vector dramatically reduces the severe germline loss phenotype produced when this

construct is in the VALIUM 20 vector. Similarly, cloning the VALIUM 22 GL00188 shmiRNA into VALIUM 20 increases the severity of its germline loss phenotype (table 4.2, table 4.3). This set of experiments not only demonstrate the importance of vector choice but also the importance of choosing high efficacy shmiRNA constructs when knocking down gene function by RNAi. This weakness of RNA interference has been well documented in many model organisms, and it poses similar challenges for the analysis of spliceform contributions to gene function. A weak spliceform phenotype could mean that that particular spliceform is not as important as other spliceforms of a gene, or that RNAi knockdown of that spliceform is simply incomplete. The recent advancements of the CRISPR-cas9 genome editing system, makes it possible to do a more thorough study of spliceform contributions.

But this raises another caveat that we encounter when analyzing gonad RNAseq data. The gonads are made up of two main types of cells: germ cells and somatic cells. Therefore, RNA-Seq data collected from gonads are always a mix of expression from both types of cells. While *bam* mutant gonads do enrich for undifferentiated germ cells, there are still many somatic cells in these gonads. So many of the splicing differences selected from this data could be a result of somatic expression. A good example of this is the expression of the dlg apparent female-specific isoforms, which are in fact expressed in the soma of both sexes

(Fig4 A-B, F). Analyzing the expression pattern of isoform-specific expression constructs for the other gene candidates would tell us which cell type they are expressed in. Additionally, using the RNAi constructs from this study to knock down gene expression in the soma will tell us whether any of these genes have important functions in the somatic cells of the gonad.

Since the conclusion of this project, I have conducted another RNA-Seq experiment using *bam* mutant male and female gonads. Improvements to sequencing technology since Q. Gan's experiment allowed for deeper sequencing of the gene expression in these gonads. Surprisingly, only half of the genes selected in this study also have differential exon expression in my bam RNA-Seq data. It is possible that the increased sequence depth resulted in the loss of many of the subtle splicing differences seen in Q. Gan's data. Importantly, most of the previously characterized sex-specific splicing events can be seen in both datasets (Fig 4.5 A-B), though their fold changes would not all meet the requirements used for selection in this project. It would be interesting to see the outcome of re-doing this project using a context driven selection method instead of a fold change driven selection method.

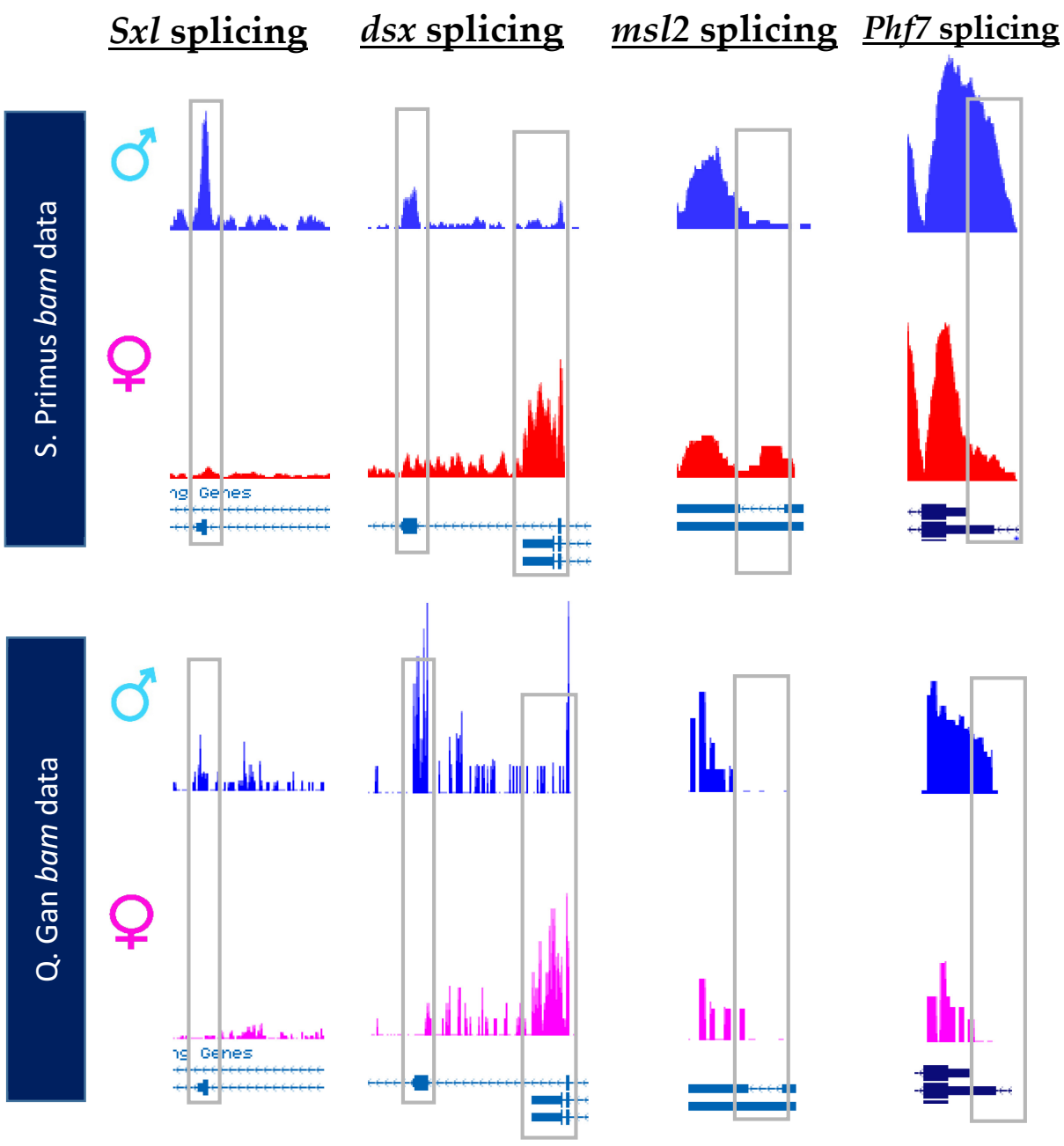



**Figure 4.5: RNA-Seq data showing sex-specific splicing of known genes.** A) Splicing of *Sxl*, *dsx*, *msl2*, and *Phf7* in my *bam* mutant RNA-Seq data (unpublished). B) Splicing of *Sxl*, *dsx*, *msl2*, and *Phf7* in Gan et al 2011 *bam* mutant RNA-Seq data.

# Chapter 5:

# *dlg* is required for proper development of the male germline

## Introduction:

Gametogenesis is a complex process that requires specific inputs to produce the right output. Oogenesis and spermatogenesis have many unique features; the biggest of which is their production of very distinct sex-specific gametes; eggs and sperm. However, the core steps in oogenesis are very similar to those in spermatogenesis: they both begin with the asymmetric division of germline stem cells (GSCs), followed by four incomplete mitotic divisions, and finally entry into meiosis and terminal differentiation of the gametes. Owing to these similarities, the mechanisms used to complete some of these processes are shared between the sexes. For example, the *bag of marbles* (*bam*) gene is essential to initiate germline differentiation in both sexes (McKearin and Spradling (1990), and both sexes require hu li tai shoa (hts) for proper formation of the ring canals which connect germ cells of a single cyst (Yue and Spradling, 1992; McKearin and Ohlstein, 1995). Despite these similarities, however, many processes are controlled using sex-specific mechanisms. For example, maintaining stemness in germline stem cells is integral in both sexes for the continued production of gametes. But the main mechanism used to maintain GSCs is different between the sexes—males primarily use the *Jak*/*Stat* pathway while females use the *TGFb* pathway.

The function of many genes have been studied to understand their sex-specific roles in gametogenesis. The screen described in chapter 4 discovered several such genes. While this screen was undertaken primarily to discover spliceforms with sex-specific roles in gametogenesis, it was designed in such a way to also discover genes with sex-specific roles in gametogenesis. One very interesting gene to come out of this screen was a gene called *discs large* (*dlg*), a membrane-associated protein localized to septate and neuromuscular junctions in Drosophila, complexed with two other proteins *scribble* (*scrib*) and *lethal giant larvae* (*lgl*). The *dlg* gene is well studied in epithelial cells of flies, worms and mammals, and has been found to perform many important functions. In short, it plays important roles in maintaining apicobasal cell polarity, as well as regulating cell proliferation, survival, migration and differentiation (reviewed in Humbert et al, 2008). There have also been studies of this gene in the ovarian soma (Bilder et al, 2000) and in the somatic cells of the testis (papagiannouli & Mechler, 2009). But a study of its role in the germline has not been possible until recently for one main reason: RNA interference has not been strong enough to produce a germline phenotype. The advent of the VALIUM 20 and 22 vector RNAi lines (Ni et al, 2011) with their superior expression in the Drosophila germline, has unmasked dlg's role in the germline as separate and distinct from its somatic function.

*dlg* belongs to the family of membrane-associated guanylate kinases (MAGUKs).This family of proteins is characterized by 1 to 3 PDZ domains (PSD-95, Dlg, ZO-1), an SRC homology 3 domain (SH3), a HOOK domain, and a GUK domain, which is highly similar to guanylate kinase (Woods and Bryant, 1991; Woods et al, 1996). The specific contributions of these domains to *dlg's* different functions has been studied in epithelial cells (Hough et al, 1997). *dlg* is frequently referred to as a tumor suppressor gene, because the phenotype caused by its loss in epithelial cells is overgrowth and overproliferation resulting in tumorous tissues. The tumor formation is thought to come about in part by the loss of epithelial cell polarity. The function of *dlg* in maintaining cell polarity made it initially very interesting to us as a good candidate for regulating the highly polarized divisions carried out by germline stem cells. My data provides evidence supporting the hypothesis that *dlg* plays some small role in GSC polarity. In this chapter, I discuss my characterization of the male-specific phenotype caused my *dlg* loss from the germline. My investigation finds that *dlg* is required male-specifically in the Drosophila germline for proper germline development. Unlike in epithelial cells, its loss in the germline does not result in tumor formation, instead, *dlg* mutant germ cells cease to proliferate and are lost from the gonad. The germ cell loss occurs via a *debcl*-mediated cell death mechanism, which can be thwarted by loss of *debcl* function, partially rescuing

the germ cell loss phenotype. Genetic interactions between *dlg* and *Ecadherin* (a key component of adherens junctions), as well as *zpg* (a key gap junction protein) also suggests that *dlg*'s function in the germline is important for maintaining junctional integrity in male germ cells.

## Materials and Methods

### Fly stocks

The fly stocks used were obtained from Bloomington Stock Center unless otherwise indicated. Nanos-Gal4 (Van Doren, 1998), $zpg^{z\text{-}2533}$, $zpg^{z\text{-}5352}$, uas-Ecad::GFP

### Immunofluorescence and centrosome localization assay

Adult ovaries and testes were fixed, blocked and stained as described in Gonczy et al, 1997. All images were taken with a Zeiss LSM 510 confocal microscope.

Primary antibodies and the concentrations used are as follows: chicken anti-Vasa 1:10000 (K. Howard); rabbit anti-vasa 1:10,000 (R. Lehmann); rat anti-Ncadherin 1:12 (DN-EX#8, DSHB); mouse anti-HTS 1:4 (1B1, DSHB); mouse anti-armadillo 1:100 (N2 7A1, DSHB); mouse anti-dlg1 1:20 (4F3, DSHB); mouse anti-y-tubulin 1:100 (Sigma); rabbit anti PH3 1:5000 (Millipore); rabbit anti-GFP 1:1000 (1:1000);

rabbit anti-activated caspase 3 1:100 (BD Pharminogen); mouse anti-Fas3 1:30 (7G10, ,DSHB); rabbit anti-Stat92E (Montell lab). DSHB: Developmental Studies Hybridoma Bank. Secondary antibodies were used at 1:500 (Alexa-fluor). Stains were mounted in vectashield mounting solution with DAPI (vector Industries).

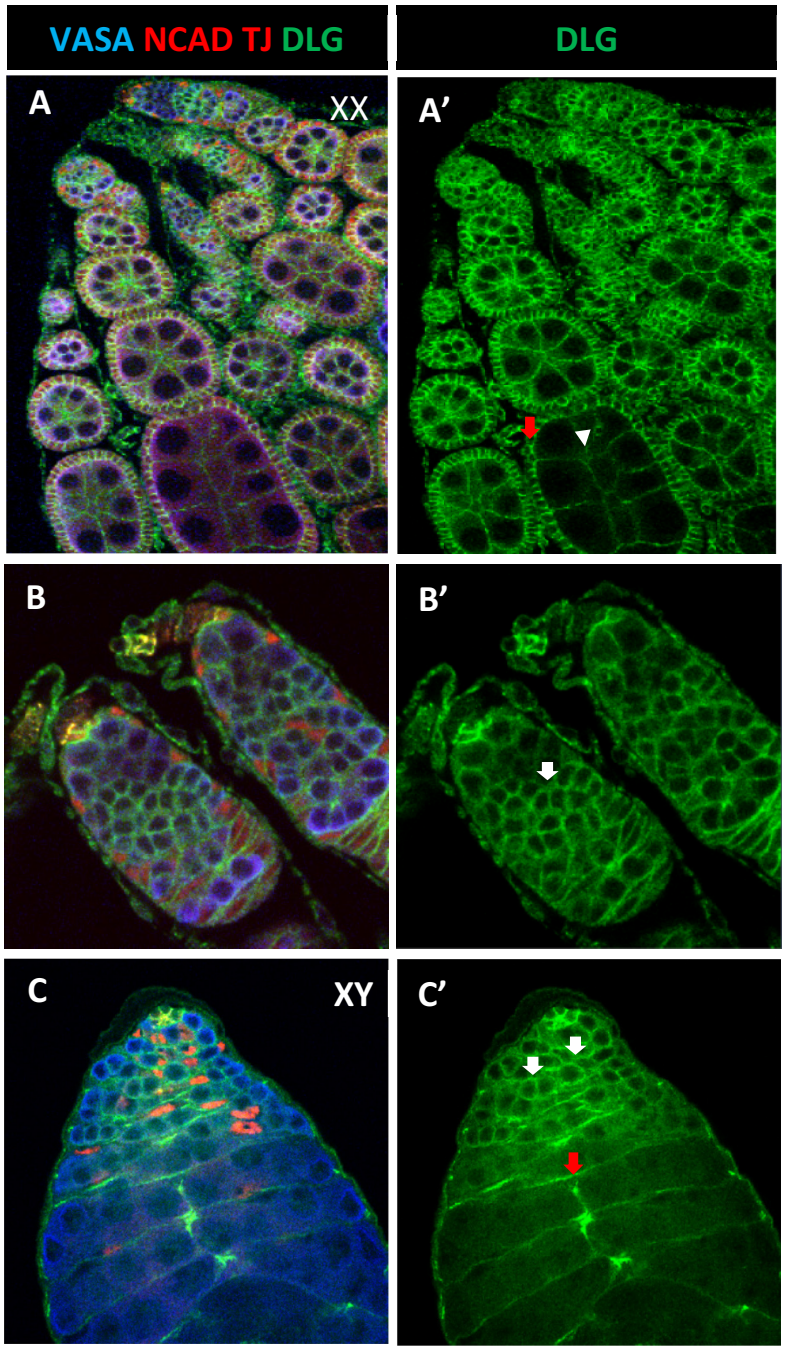
GSCs were scored for centrosome localization using y-tubulin to mark centrosomes. Only GSCs with more than one centrosome were scored. GSCs with at least one centrosome at the hub–GSC interface are defined as oriented and those without a centrosome at the interface are misoriented. % misoriented centrosomes = # of misoriented centrosomes/ Total # of GSCs scored.

## Results

### Dlg is needed cell-autonomously in the male germline

DLG protein is expressed in the gonads of both sexes, localized to germ cell and somatic cell membranes (Fig 5.1). At the anterior of the gonads where gametogenesis begins, dlg is expressed in all cells (Fig 5.1B′, C′ white arrows). Later in gametogenesis, however, dlg expression becomes restricted to somatic cells in males (Fig 5.1C′ red arrow).

**Fig 5.1: Expression of dlg in ovaries and testes**. A-B) Confocal images of adult ovaries. B) germaria showing dlg expression in all early germ cells and somatic cells. C) Confocal image of an adult testis showing dlg expression in early germ cells of 2 and 4 cell cysts, but absent from germ cells of later cysts. Red arrows point to somatic dlg expression. White arrows point to dlg expression in the early germline. White arrowhead shows dlg expression in the late germline in females only.


VASA NCAD TJ DLG
DLG
A
XX
A'
B
B'
C
XY
C'

To determine whether dlg was necessary for proper germline function I knocked it down in the germline by RNA interference. Loss of dlg from the germline causes a very dramatic male-specific phenotype. Males without dlg function in the germline undergo germline depletion and adults have completely lost their germline by eclosion (Fig 5.2A-B). Somatic cells of the gonad are still present in dlg-RNAi testes (Fig 5.2B'' arrowhead). This sex-specific germline loss phenotype was confirmed using several different TRiP RNAi lines as well as an RNAi line that I created (Chapter 4: table 4.2 and 4.3). Females without dlg function in the germline show no ovary defect (Fig 5.2 C-D) in most RNAi lines used except for the $dlg^{d8\text{-}g2}$ RNAi line which caused an egg chamber development defect (data not shown) but no germline loss was observed.

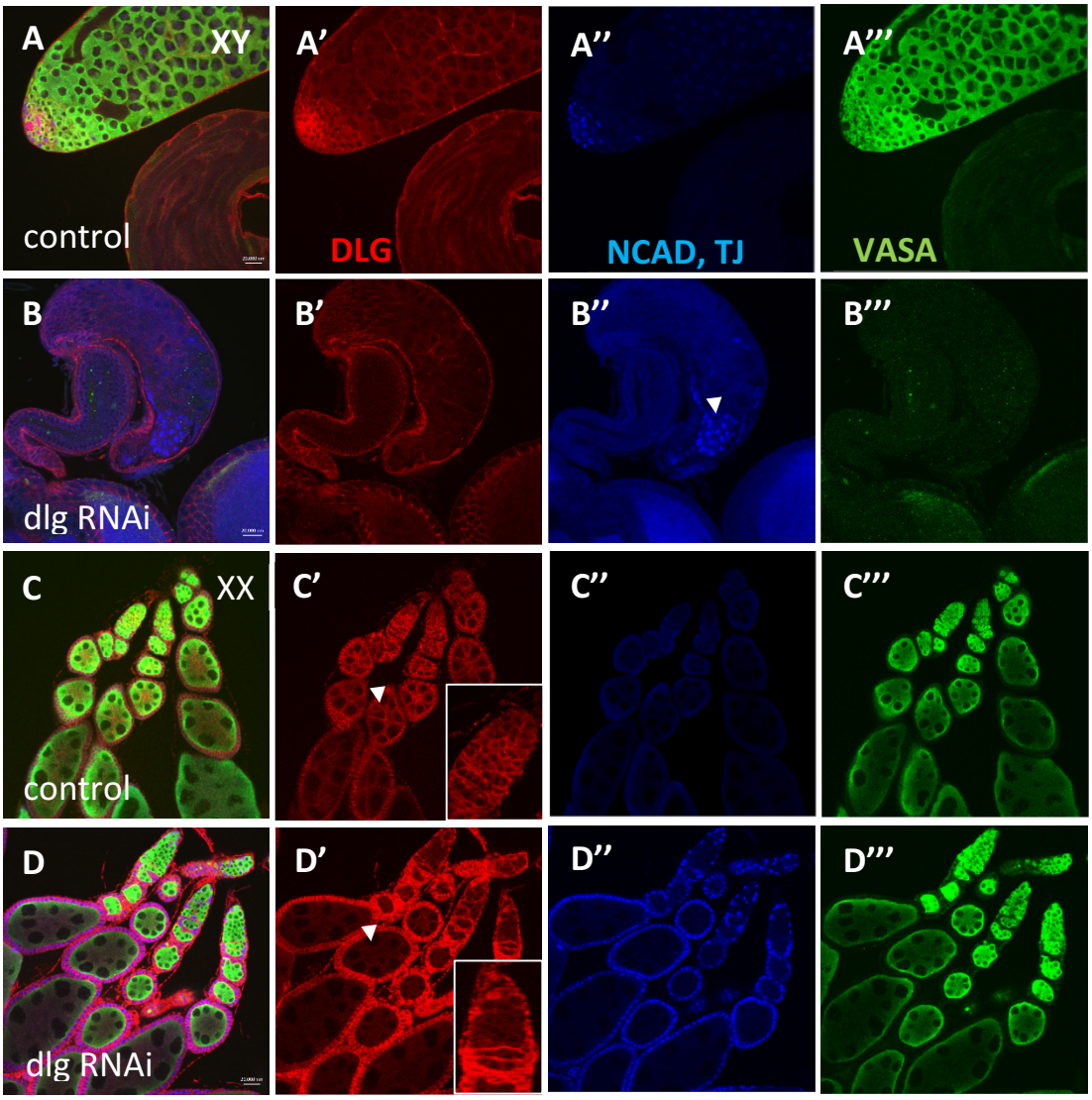
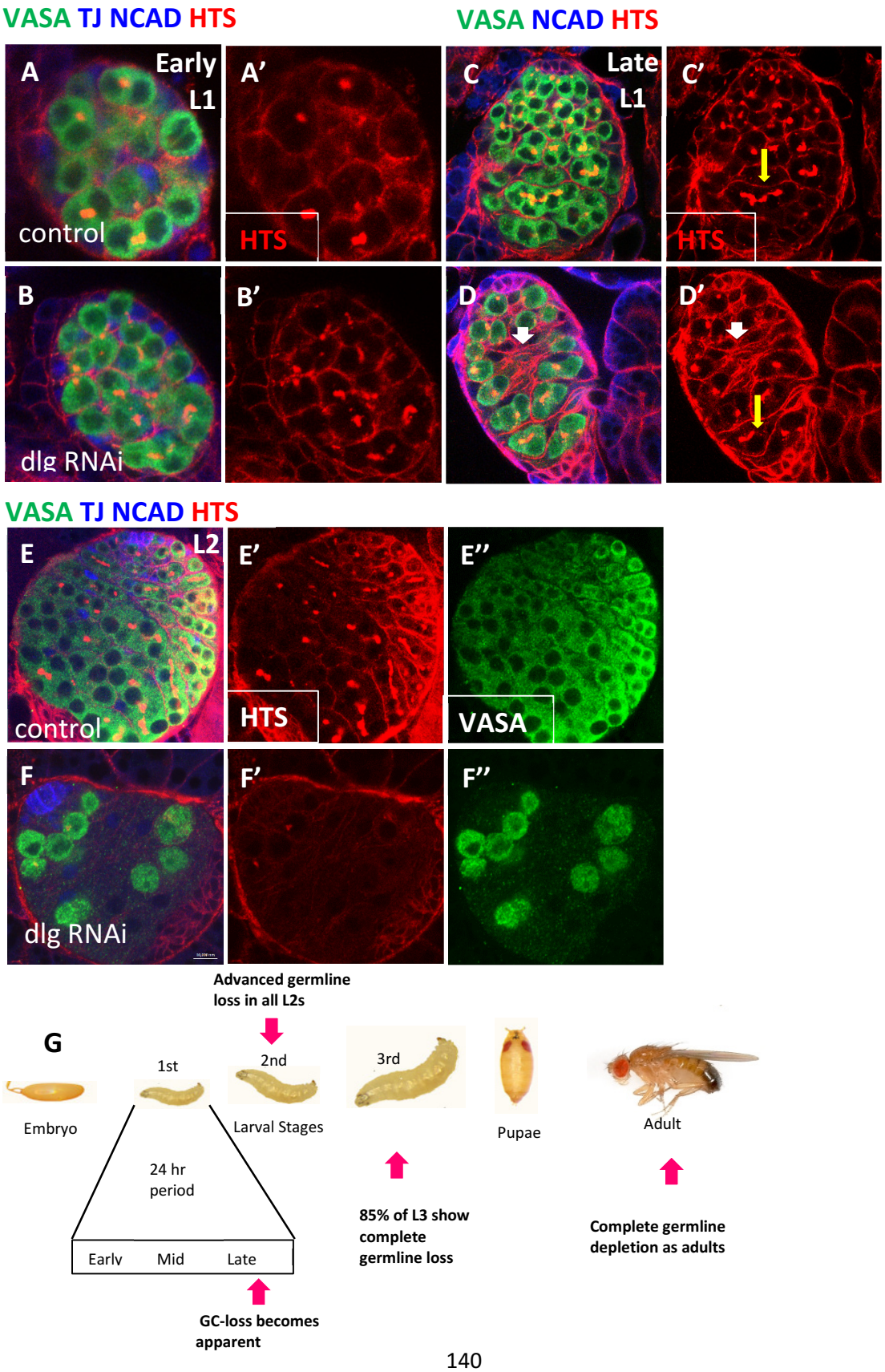



**Fig 5.2: Loss of dlg from the germline results in male-specific germline depletion**. A-B) Confocal images of A) nanos-Gal4 control and B) nanos-Gal4 > *dlg*-RNAi testes. Arrowhead marks the somatic cells remaining in the gonad. C-D) Confocal images of C) nanos-Gal4 control and D) nanos-Gal4 > *dlg*-RNAi ovaries. Insets are closeups of the germaria showing the absence of dlg in the germline of the RNAi compared to the presence of dlg in control ovaries. Arrowheads show the same loss in the late germline.

To characterize the male germline loss phenotype it was necessary to analyze dlg-RNAi gonads that still had germline remaining. Throughout this project I did this in 3 different ways: I used weaker RNAi lines that result in only mild to moderate germline loss in adult males, I expressed a strong dlg-RNAi using zygotic nanos-Gal4, which produces a milder phenotype compared to the use of maternal and zygotic nanos-Gal4, and I used the larval gonads of animals with strong dlg-RNAi knockdown in the germline. Examination of larval gonads revealed that while early L1 larval gonads all look normal, in late L1 larval gonads germline loss becomes apparent (Fig 5.3A-D white arrow). Germ cell loss can be seen both in germline stem cells as well as in developing cysts. In addition, while there are many 8 cell cysts and rare 16 cell cysts in wild type late L1 gonads, dlg-RNAi gonads rarely have 8-cell cysts. This suggests that there is a delay in cyst development of dlg-RNAi gonads. By larval stage 2 most dlg-RNAi gonads have very few germ cells remaining (Fig 5.3E-G).

**Fig 5.3: Germline loss in dlg-RNAi testes become apparent in late L1 larval gonads.** A-B) Confocal images of A) nanos-Gal4 control and B) nanos-Gal4 > *dlg*-RNAi early L1 testes. C-D) Confocal images of C) nanos-Gal4 control and D) nanos-Gal4 > *dlg*-RNAi Late L1 testes. Germline loss is evident in RNAi testes (white arrows). Yellow arrows show the oldest cyst in each gonad. E-F) Confocal images of E) nanos-Gal4 control and D) nanos-Gal4 > *dlg*-RNAi Larval L2 testes showing advanced germline loss in dlg-RNAi testes. G) Drosophila life cycle timeline showing the progression of the germline loss phenotype


VASA TJ NCAD HTS
VASA NCAD HTS
A
Early L1
A'
control
HTS
C
Late L1
C'
HTS
B
B'
dlg RNAi
D
D'
VASA TJ NCAD HTS
E
L2
E'
E''
control
HTS
VASA
F
F'
F''
dlg RNAi
Advanced germline loss in all L2s
G
1st
2nd
3rd
Embryo
Larval Stages
Pupae
Adult
24 hr period
Early
Mid
Late
85% of L3 show complete germline loss
Complete germline depletion as adults
GC-loss becomes apparent

## Germline loss occurs by debcl-mediated cell death

To examine the germ cell death further, I looked at the mechanism by which dlg-RNAi germ cells were dying. First I looked at caspase-mediated cell death to determine whether the germ cells were dying by a caspase-mediated mechanism. This is the most common form of apoptosis and can be monitored by immunostain for activated-caspase 3. The caspase-3 stain was done in adult gonads expressing *dlg*-RNAi with zygotic only *nanos*-Gal4. Adult testes provide a built-in positive control for caspase stain, because of the cytoplasmic waste that is stripped from sperm during sperm individualization. The caspase stain showed no significant expression of caspase3 in any of the remaining germ cells of these adults Fig 5.4A-B). The caspase stain seen in both control and RNAi images depict sperm individualization. Similarly a caspase stain done in late L1 gonads when loss of germ cells is already evident, shows no activated caspase expression in dlg-RNAi germ cells (data not shown). Caspase mediated cell death can be blocked by the expression of the baculovirus caspase inhibitor P35. Ectopic expression of P35 in dlg-RNAi germ cells did not rescue the germ cell death phenotype (Fig 5.4C). All this evidence suggests that the germ cells are dying by a non caspase-mediated mechanism.

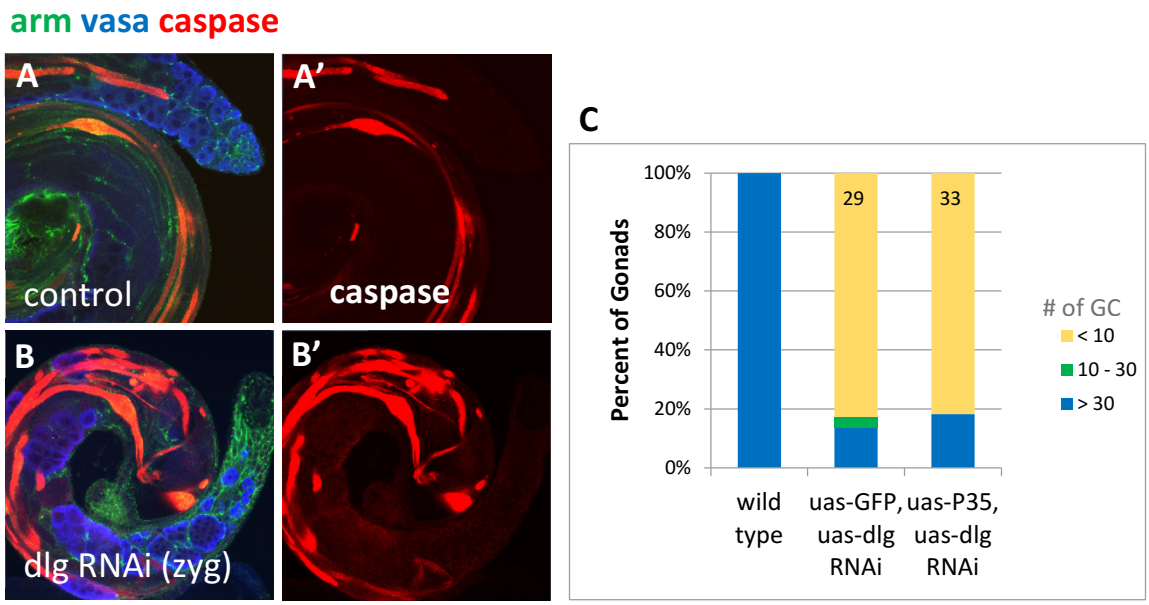


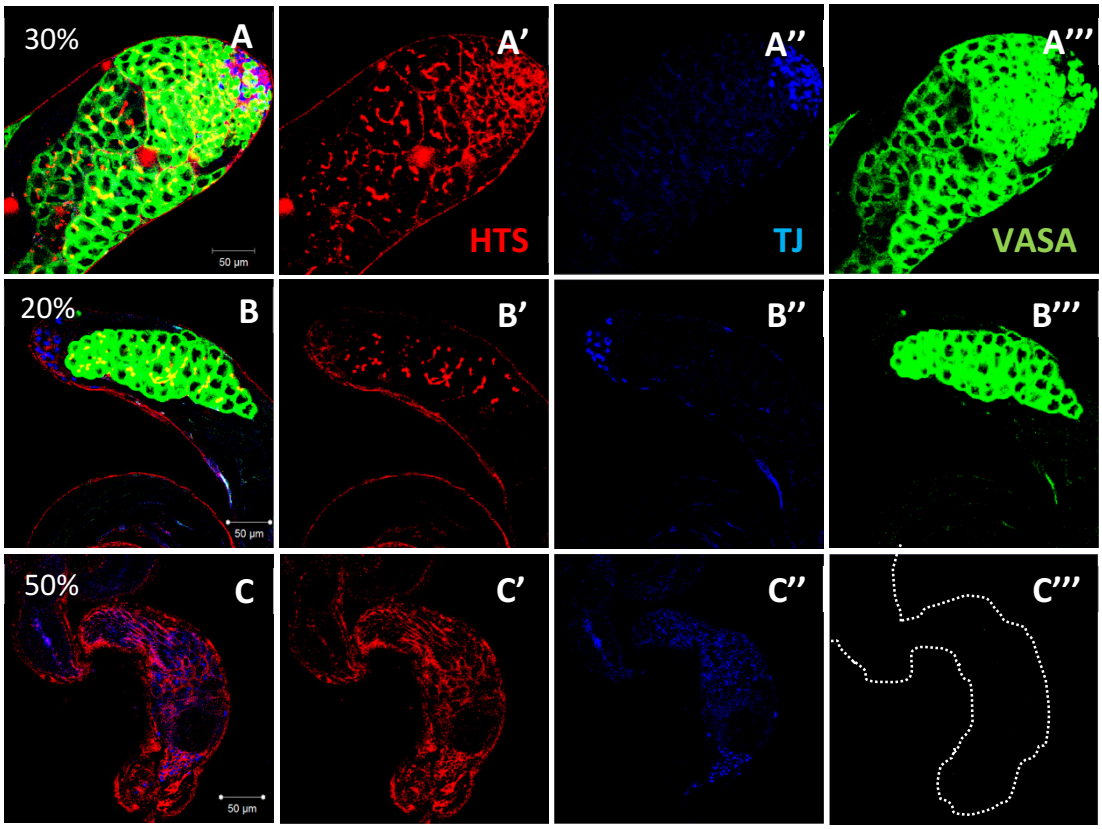

**Fig 5.4: germ cells lacking dlg are dying by a NON caspase-mediated cell death mechanism**. A-B) Confocal images showing A) zygotic nanos-Gal4 control and B) zygotic nanos-Gal4 > *dlg*-RNAi adult testes. Germline loss phenotype is weaker because maternal gal4 is not present. Caspase stain can be seen in individualizing sperm but not in germ cells. C) Graph showing quantification that inhibition of caspase-mediated cell death by ectopic expression of P35 does not rescue germ cell death in dlg-RNAi testes.

The finding that Germ cells frequently die by a non-caspase mediated mechanism is well-known in the field. Recently it was shown that developmental germ cell death (GCD), where "male germ cells are spontaneously eliminated before entering meiosis" (Yacobi-Sharon et al, 2013), occurs by an alternative cell death pathway involving the Drosophila Bcl-2 family proteins (Yacobi-Sharon et al, 2013). Debcl is one of the proteins in this family. I investigated whether this alternative cell death pathway was involved in the death of germ cells lacking dlg by knocking down debcl function in germ cells also lacking dlg. This produced a dramatic rescue of the dlg-RNAi germline loss phenotype. Thirty percent of debcl-rescued testes appear wild type-like in testis size (Fig 5.5D-E), as well as in number and distribution of germ cells (Fig 5.5A-A'''). Germ cells form cysts and go on to differentiate into sperm that is wild type in appearance (Fig 5.5F-G). Sperm from debcl-rescued testes is also functional as these animals are fertile. This suggests that germ cell death contributes greatly to the germ cell loss phenotype. But germ cell death is not the only contributor to germ cell loss. Twenty percent of debcl-rescued testes have more germ cells than a typical dlg-RNAi but many stages of germline differentiation are missing (Fig 5.5B-B'''). And another 50% of debcl-rescued testes show no rescue of the germ cell loss phenotype (Fig 5.5C-C'''). This group includes some testes that are wild type in testis size but have already lost all of their germline. In addition, testes from

debcl-rescued animals become more and more germ cell-less as they age. By 15 days old, debcl-rescued testes have lost all of their germ cells.

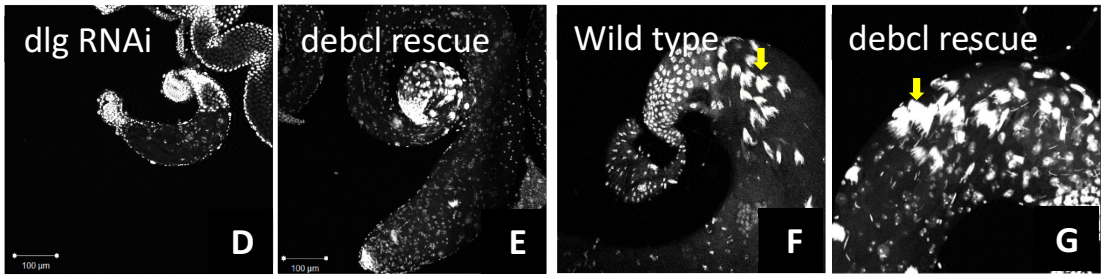
**Fig 5.5: knockdown of debcl rescues germ cell death, testis size and sperm production.**. A-B) Confocal images showing the degrees of rescue in 0-2 day old debcl rescue testes. Genotype is nanos-Gal4; dlg-RNAi/debcl-RNAi. Debcl rescue n=89 compared to nanos-Gal4;uasGFP control n=65. A-A′′′) Best rescued testes are full of germline and appear wild type-like. B-B′′′) Minimal rescued testes have between 10 and 30 germ cells C-C′′′) Confocal images showing NON-rescued testes that look like typical dlg-RNAi testes (in fig 5.2B). D-G) Confocal images of dapi stain to show only DNA. D-E) size of the D) dlg-RNAi and E) debcl-rescued testes taken at 20X magnification. F) Sperm in wild type testes and G) debcl-rescued testes.

debcl rescue




Dapi: size

Dapi: sperm



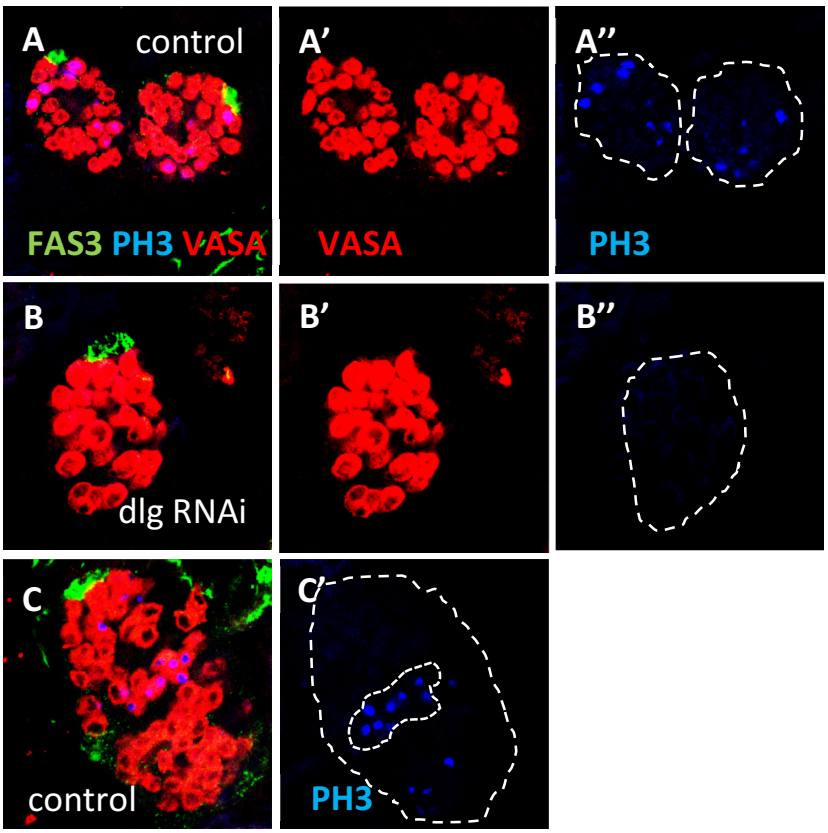
## Germ cells lacking dlg have reduced proliferation

Another observation made in Late L1 dlg-RNAi larval gonads is that there appears to be a delay in cyst development. While control larval gonads have started differentiating into 16-cell cysts at this stage, dlg-RNAi gonads have not (Fig 5.3C', D' yellow arrows). In fact at this stage, even 8-cell cysts are rare in dlg-RNAi gonads. This suggested that dlg-RNAi germ cells were not proliferating at the same rate as wild type germ cells. To test this hypothesis, I quantified frequency of cell division in control and dlg-RNAi larval gonads. Monitoring mitosis with anti-phosphorylated (ser10) histone 3 (PH3) immunostain, I found that early L1 dlg-RNAi gonads have dividing germ cells only half as much as controls do (Fig 5.6D). At this stage in development, the germ cell loss phenotype is not yet evident. This data suggests that the reduction in germ cell division precedes debcl-mediated cell death.

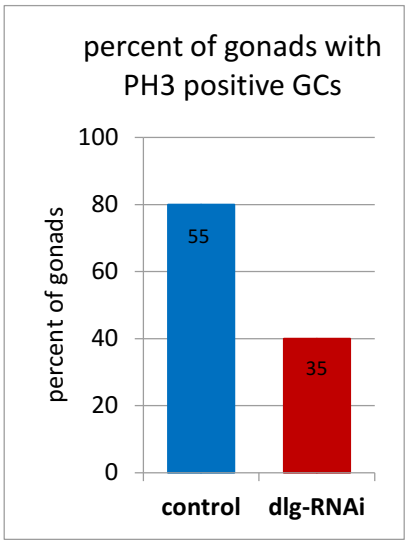
When germ cell death becomes evident in late L1, reduced mitotic division is still evident even in gonads showing the least germ cell death (Fig 5.6A-B''). This lends further support to the conclusion that reduced mitosis is not the result of having less cells but instead it precedes the loss of germ cells. One hallmark of germ cell cyst development is the synchronous division of the germ cells in the same cyst (Fig 5.6C-C'). By late L1, synchronous cyst divisions happens rarely in strong dlg-RNAi gonads, indicating that they cannot maintain

the germline differentiation program. In addition to a reduction of cyst divisions, the mitotic index of dlg-RNAi germline stem cells is 10 fold less than control germline stem cells in late L1 larval gonads (Fig 5.6E). This evidence suggests that reduced germ cell division is a big factor contributing to the dlg-RNAi phenotype. It is possible that this cell cycle arrest is what triggers debcl-mediated germ cell death.

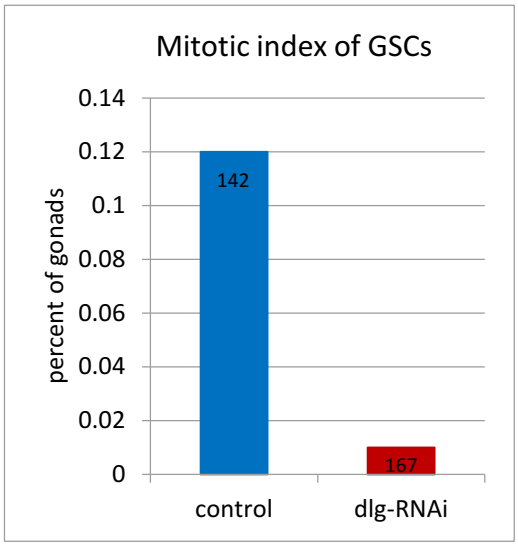
**Fig 5.6: dlg RNAi germ cells are less mitotically active**. A-B) Confocal images of A) nanos-Gal4 control and B) nanos-Gal4 > *dlg*-RNAi late L1 larval testes showing PH3 positive germline stem cells as well as other germ cells in control but no PH3 positive germ cells in dlg-RNAi. C-C′) Confocal image of a nanos-Gal4 control late L1 testis showing an 8-cell cyst undergoing synchronous division. This was never seen in testes with strong dlg-RNAi. D) Quantification of the percent of early L1 larval gonads with any PH3 positive germ cells. E) Mitotic Index of germline stem cells (GSCs) in late L1 larval gonads.


A
control
FAS3 PH3 VASA
A'
VASA
A''
PH3
B
dlg RNAi
B'
B''
C
control
C'
PH3

D


percent of gonads with PH3 positive GCs
percent of gonads
100
80
60
40
20
0
55
35
control
dlg-RNAi

E


Mitotic index of GSCs
percent of gonads
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
142
167
control
dlg-RNAi

## Some niche–GSC interactions are affected in germ cells lacking dlg

Germline stem cells are a unique kind of germ cell. They begin gametogenesis and continually divide to maintain gamete production. GSCs divide asymmetrically, producing one daughter cell that goes on to differentiate into gametes while the other daughter cell retains its stem cell identity. The ability of a GSC to maintain stem cell identity is mediated by signaling from the GSC niche, a microenvironment which nurtures and maintains stem cells. The observation that GSCs are equally affected by the loss of dlg from the germline, suggests niche-GSC interactions may also be defective in germ cells lacking dlg. The major signaling pathway used to maintain GSCs in the male niche is *Jak*/*Stat* signaling. Proper *stat* signaling is important to maintain *E-cadherin*-mediated GSC attachment to the hub (Leatherman and Dinardo, 2010). To investigate niche–GSC interactions I first examined whether GSCS lacking dlg can respond to Jak/stat signaling. I found that GSCs in early L1 larvae are able to properly upregulate stat in response to signaling from the niche (Fig 5.7).

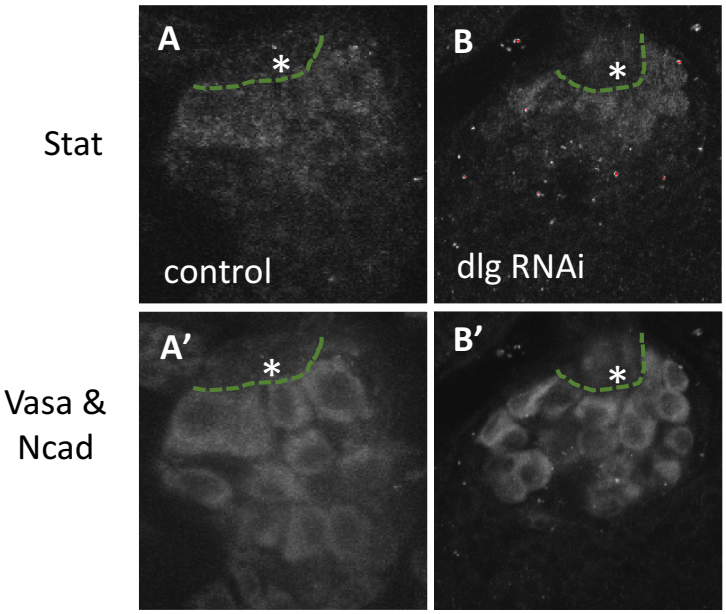
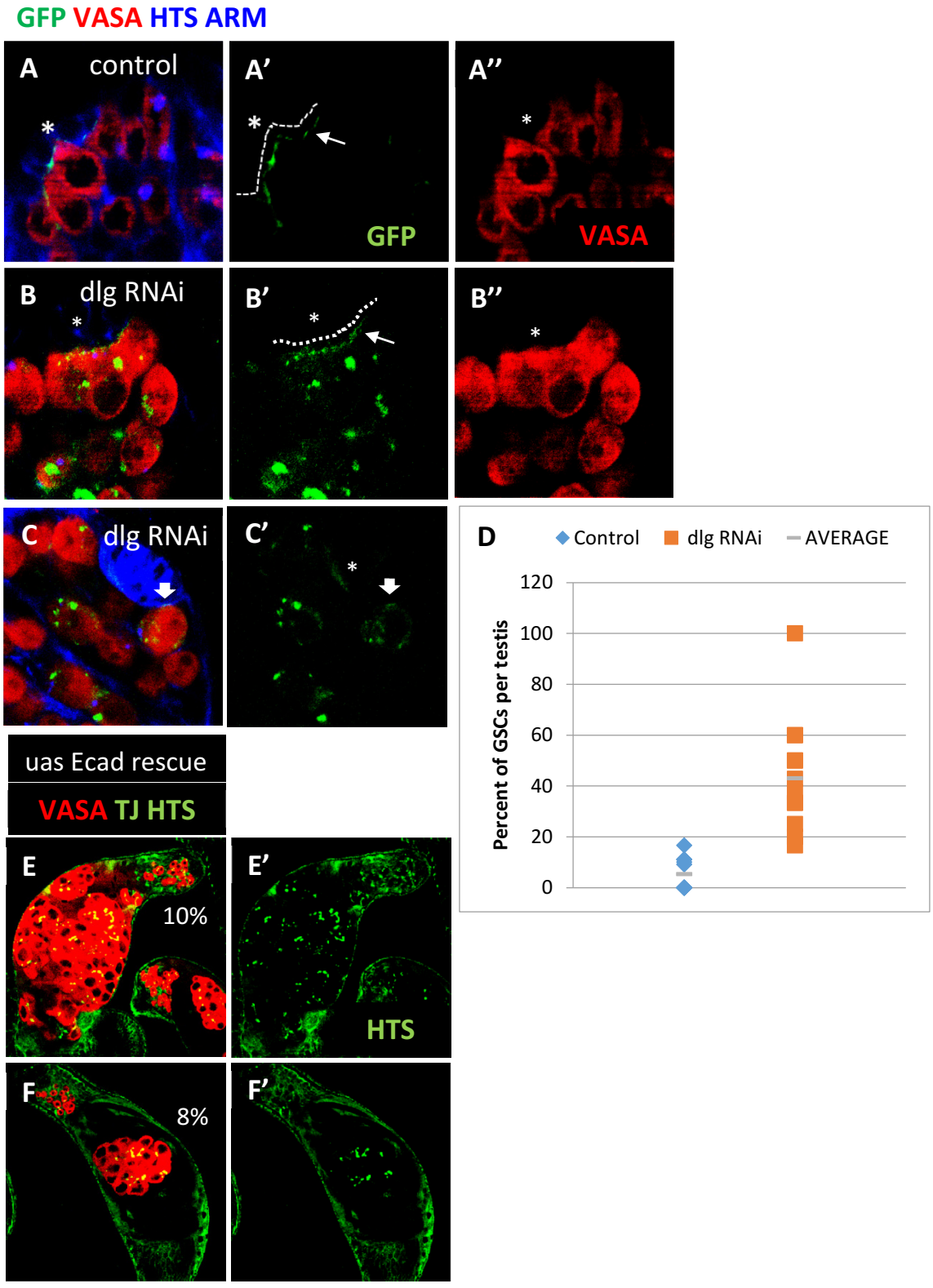



**Fig 5.7: GSCs lacking dlg can respond to jak/stat signaling from the niche.**

Confocal images of A) nanos-Gal4 control and B) nanos-Gal4 > *dlg*-RNAi early L1 testes showing stat stain in germline stem cells.

Proper response to stat signaling suggests that *E-cadherin*-mediated GSC attachment to the hub is also properly functional in GSCs lacking dlg. To investigate whether this was true, I ectopically expressed an Ecad::GFP construct in germ cells and assayed for GFP localization to the hub–GSC interface. While Ecad::GFP localized properly to the hub–GSC interface (Fig 5.8A′, B′ arrows), much of the GFP also accumulated into cytoplasmic punctae in the dlg-RNAi germ cells (Fig 5.8B′, C′, D). This suggests that trafficking of Ecad may be defective in germ cells lacking dlg. Another common observation with Ecad::GFP localization at the hub–GSC interface is a decreased area with which remaining GSCs are attached to the hub (Fig 5.8C-C′ arrow). It is possible that loss of dlg causes a change in GSC cell adhesive properties resulting in GSCs releasing their attachment to the hub, consequently losing their stemness and turning on the differentiation program. GSCs that leave the niche differentiate out of the testis thereby reducing the testis' ability to replenish the germline. This occurrence will contribute to the overall germline loss phenotype caused by loss of dlg. Over-expressing a wild type copy of E-cadherin in the germline resulted in some rescue of the dlg RNAi phenotype (Fig 5.8E-F). This Ecad overexpression rescue at about 18%, though less robust than the rescue that results from inhibiting cell death, it suggests that maintaining GSC attachment to the niche is an important measure to protect against the germline loss phenotype caused by loss of dlg.

**Fig 5.8: Ectopic expression of Ecad results in some rescue of the dlg-RNAi phenotype.** Confocal images of A) nanos-Gal4, uas Ecad::GFP control and B) nanos-Gal4 > *dlg*-RNAi, uas Ecad::GFP late L1 testes showing localization of GFP to the hub–GSC interface as well as cytoplasmic GFP accumulation. C-C′) Confocal image of nanos-Gal4 > *dlg*-RNAi, uas Ecad::GFP late L1 testis with arrow pointing to Impairment of GSC adhesion to the hub. Asterisk marks the hub. Dotted line marks the hub–GSC interface. D) Quantification of the cytoplasmic GFP punctae accumulation. Plotted is the percent of GSCs with cytoplasmic EcadGFP punctae per testis. Horizontal line is the average of all the testes. E-F′) Confocal images showing the degrees of rescue in dlg-RNAi testes with ectopic expression of E-cadherin.


GFP VASA HTS ARM
A control
A'
A''
GFP
VASA
B dlg RNAi
B'
B''
C dlg RNAi
C'
uas Ecad rescue
VASA TJ HTS
E
10%
E'
HTS
F
8%
F'
D
Control
dlg RNAi
AVERAGE
Percent of GSCs per testis
120
100
80
60
40
20
0

## Centrosome orientation is affected in GSCs lacking dlg

Many studies have shown that dlg plays a major role in maintaining polarity in epithelial cells as well as in neuroblasts (Tanentzapf & Tepass, 2003; Siegrist & Doe, 2005, reviewed in Humbert et al, 2007). Maintaining polarity is also essential for germline stem cells to function properly. GSC divisions are highly polar, and are guided by centrosomes orienting themselves perpendicular to the hub. When GSC centrosomes are misoriented, this engages the centrosome orientation checkpoint (COC), effectively arresting the cell cycle for some time so that centrosomes can be properly oriented (Cheng et al, 2008). The decrease in the mitotic index of GSCs lacking dlg could be due in part to misoriented centrosomes activating the COC. To test this possibility, I analyzed centrosome orientation in early L1 and late L1 larval gonads. Indeed, centrosomes were 3 times more misoriented in Late L1 germ cells lacking dlg (Fig 5-9A-D). Centrosome orientation was similar to control levels in early L1 but became more mis-orientated in Late L1 as the germline loss phenotype also became evident. This data suggests that dlg is required for proper GSC centrosome orientation. Interestingly, in addition to misoriented centrosomes, one or more ectopic centrosomes were present in 20% of the dlg-RNAi GSCs scored (Fig 5.9C blue arrowhead), implicating dlg in regulating centrosome number as well.

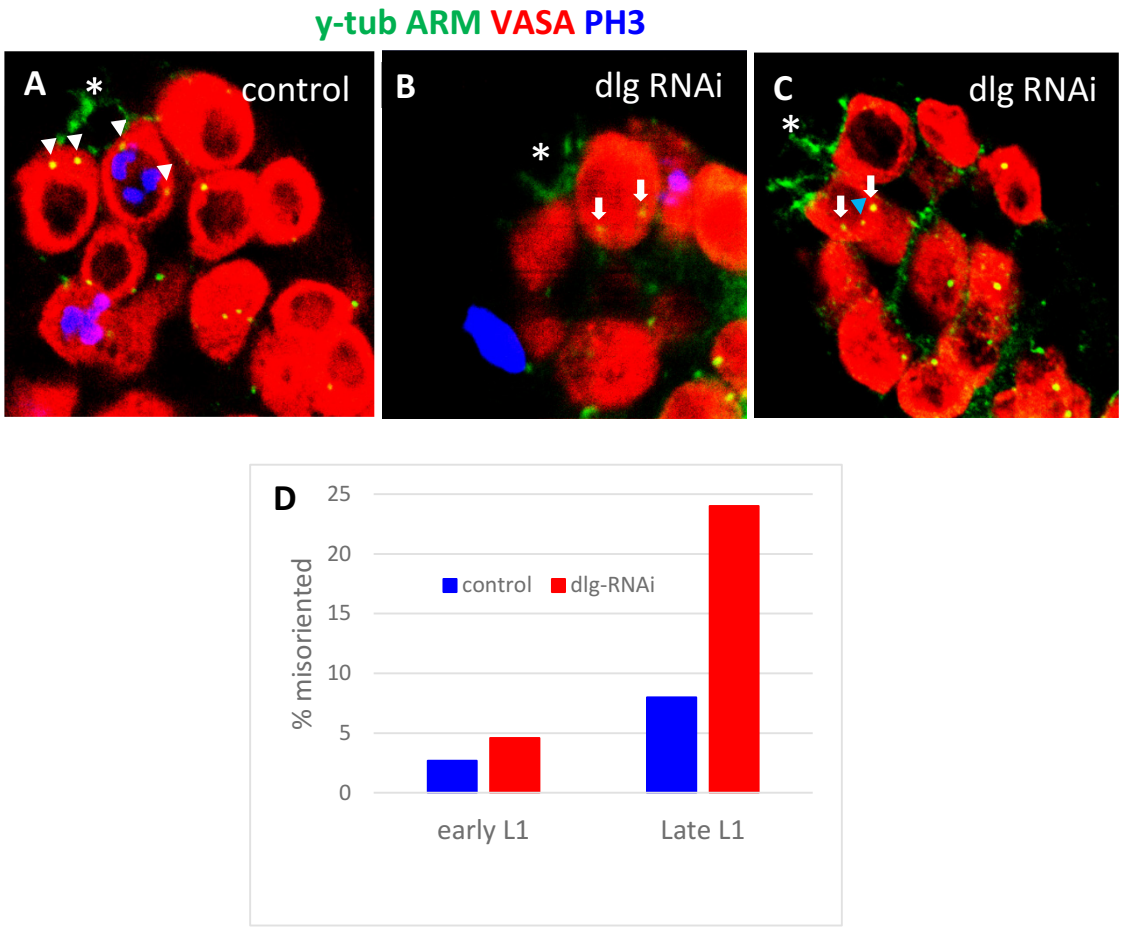



**Fig 5.9: GSCs lacking dlg have a higher percentage of mislocalized centrosomes.** Confocal images of A) nanos-Gal4 control and B-C) nanos-Gal4 > *dlg*-RNAi late L1 larval testes showing centrosome orientation. White arrowheads mark oriented centrosomes. Arrows mark misoriented centrosomes. Blue arrowhead marks an extra centrosome. Asterisk marks the hub. D) Quantification of misoriented centrosomes in early and late L1 larvae.

*dlg* genetically interacts with the gap junction gene *zpg*

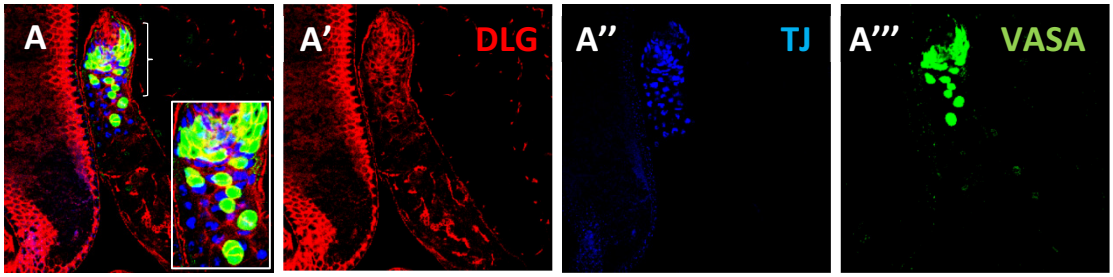
In addition to septate junctions and adherens junctions, gap junctions are another type of junction that play an important role in the germline. The gene *zero population growth* (*zpg*) is a germ cell specific gap junction gene that is required for survival of differentiating germ cells. *zpg* mutant gonads are very small and almost devoid of germ cells except for a few early germ cells (Tazuke et al, 2002). While *zpg* is needed in both sexes, it's localization has a very sex-specific pattern. In females it is highly expressed in all early germ cells including germline stem cells, while in males it is lowly expressed in GSCs (Smendziuk et al, 2015) and expressed only in small patches in early spermatogonia, then becomes highly expressed in later spermatogonia and spermatocytes evenly across the germ cell surface. (Tazuke et al, 2002). In contrast to ZPG expression, DLG is expressed strongly on the cell surface of all early germ cells up to 8-cell spermatogonial cysts, but is not expressed in germ cells past this point (Fig 5.1C'). The striking antithesis in the expression pattern of these two junction proteins implied a biological significance and suggested that they may have a mutually antagonistic relationship.

To investigate this possibility, I first examined DLG expression in a *zpg* mutant and as expected the few germ cells still present in a *zpg* mutant all express DLG (Fig 5.10A-A'''). This confirms what was reported by Tazuke et al,

that the germ cells remaining in a *zpg* mutant are GSC-like cells and early spermatogonia. These are also the germ cells that express dlg. To investigate the genetic interaction between these 2 genes, I repeated the dlg-RNAi experiment in the background of a zpg heterozygous mutant. Astoundingly, loss of 1 copy of *zpg* significantly rescues the *dlg* RNAi germline loss phenotype. This dramatic rescue produced almost 30% of testes that were wild type-like in testis size, germ cell quantity as well as germline differentiation (Fig 5.10B). Similar to the debcl-rescued testes, zpg-rescued testes also produced sperm (Fig 5.10D-E). Many of the rescues including those with very few germ cells had several sperm bundles. Another 23% of zpg-rescued testes showed a milder rescue of germ cell loss as well as testis size (Fig5.10C). Also similar to debcl-rescued testes, zpg-rescued testes lost germline over time (Fig 5.10F), signifying other inputs necessary for the maintenance of these germ cells. This data suggests that indeed dlg and zpg have an antagonistic relationship. It also indicates that one possible function of dlg is to regulate and/or restrict the formation of gap junctions in the germline.

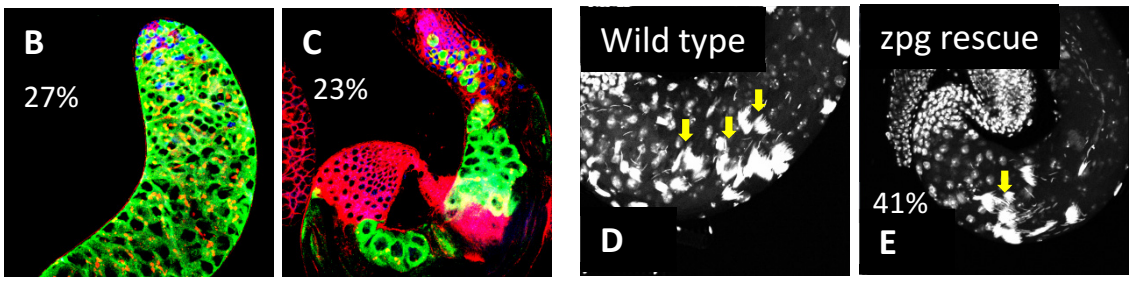
**Fig 5.10: Loss of one copy of zpg rescues the germline loss phenotype.** A-A′′′) Confocal images of *zpg* mutant adult testes. Brackets mark the remaining germline which are all *dlg*-expressing. Inset is a blowup of the bracketed region. B-C) Confocal images of adult nanos-Gal4 > *dlg*-RNAi testes rescued by the loss of one copy of zpg (zpg rescue). Percentages indicate the percent of all testes with depicted quality of germline rescue. D-E) Confocal images of adult D) wild type and E) zpg-rescued testes showing mature sperm. F) Quantification of the quality of zpg-rescue over time.

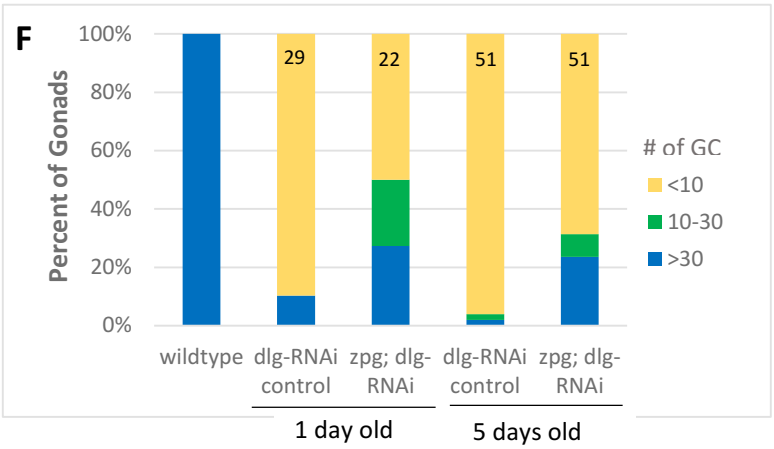
zpg mutant




zpg rescue: VASA TJ HTS ARM

Dapi: sperm



## Discussion

In this chapter, I showed that dlg is required in the male germline and affects several aspects important for proper germ cell development: germline proliferation, GSC adhesion to the hub, junctional integrity and asymmetric GSC division. The array of processes that are affected by loss of dlg is strong evidence supporting dlg as a "multitasker" in the germline, similar to the multiple roles it plays in other tissues (reviewed in Humbert et al, 2007). The debcl-mediated germ cell death that occurs in germ cells lacking dlg is likely a cell-intrinsic mechanism, activated in response to some stress signal sent by these aberrant germ cells to remove them from the testis (Fig 5.11). In fact, cell cycle arrest alone has been shown to trigger cell death. Therefore, the cumulative effects of losing dlg function in the germline is germline death. Loss of dlg from the soma also results in germline loss and testis atrophy (papagiannouli & Mechler, 2009), though its specific functions in the soma have not been explored.

The finding that dlg affects the cell cycle is not a surprise. What is surprising is the direction in which the cell cycle is changed. Many tumor suppressor genes including dlg maintain apicobasal polarity in part by negatively regulating the cell cycle. Consequently, when these genes are mutated overproliferation occurs resulting in tumors (Reviewed in Humbert et al, 2003). dlg's function in the germline, however, seems to be important for promoting

proliferation to properly regulate the germline differentiation program. When this program becomes defective, germ cell death ensues. My data shows that germ cell death is the end result of dlg loss. Indeed preventing germ cell death results in a dramatic rescue of germline loss. Preventing germ cell death may also cause a positive feedback on the germ cells prompting them to overcome the cell cycle arrest and continue proliferating. In this situation however, the germ cells especially the GSCs still are not properly maintained, and are lost from the testis by differentiating out of the testis as sperm. This may account for the age-related germline depletion observed in debcl-rescued testes.

My data also provides evidence supporting the hypothesis that *dlg* plays a role in GSC polarity. This loss of polarity in the form of 24% misoriented centrosomes (in Late L1 larval gonads) is similar to what is observed in the literature for some genes. Examples of this are E-cadherin and centrosomin, an integral centrosome component. When either gene is mutant, it results in 35% misoriented centrosomes (yamashita et al, 2003, Inaba et al, 2010). And APC and APC2 mutants have 22 to 45% and 16 to 21% centrosome misorientation respectively (Yamashita et al, 2003). However misoriented centrosomes in these mutants typically cause an increase in GSCs at the hub. Since GSCs lacking dlg are quickly lost from the niche, their loss of polarity is unlikely to be the cause of the germline loss phenotype. Therefore while dlg's role in regulating GSC

asymmetric division likely contributes to overall GSC homeostasis, this is unlikely to be the main cause of the dlg germline loss phenotype.

The relationship between dlg and E-cadherin, as well as other adherens junction proteins and other apically-localized protein complexes has been studied in Drosophila, C. elegans, and mammals (Woods et al, 1997; Bossinger et al, 2001; Firestein & Rongo, 2001; Laprise et al, 2004; Harris and Pfeifer, 2004). These complexes have been shown to have an interdependence on each other such that some are required for the proper formation of the others, and mutation or mislocalization of one causes many of the others to also become mislocalized. My data suggests that dlg and Ecad also interact in the germline but the actual effect of dlg on Ecad localization and function is not yet understood. In GSCs lacking dlg, Ecad::GFP localizes normally to the hub–GSC interface but much of the GFP is also localized in distinct cytoplasmic punctae. This could signify that dlg is important for trafficking of Ecad. Further investigation into this phenomenon is warranted. The uas-Ecad rescue is also not as robust as the $zpg^{z-2533}$ mutant or debcl-RNAi rescues. The rescue of germ cell differentiation looks more tumorous in the Ecad-rescued testes compared to the normal cyst development (monitored by HTS stain) seen in many of the debcl rescues (compare 5.8E′ to 5.5A′). In addition, the best Ecad rescue only has about 3 sperm bundles while the best zpg and debcl rescues have more than 10. Therefore,

while Ecad is essential to maintain GSCs, it is not sufficient to produce a strong rescue of the dlg-RNAi phenotype. This is likely because the defect in the germline is not restricted to germline stem cells.

Since proteins belonging to the other cell–cell junctions have all been shown to interact to some degree, it is no surprise to find that the gap junction protein *zpg* also interacts with dlg. The function of zpg was recently fully characterized in the Drosophila testis; it was found to be essential for signaling between the germline and the soma (Smendziuk et al, 2015). I propose that dlg and zpg share a similar relationship akin to the way in which the apical and basolateral complexes act antagonistically to each other, to ensure that each complex is properly localized and stays restricted to its respective functional domain. It is clear that maintaining such junctional integrity is important to maintain germ cell homeostasis and support germline development

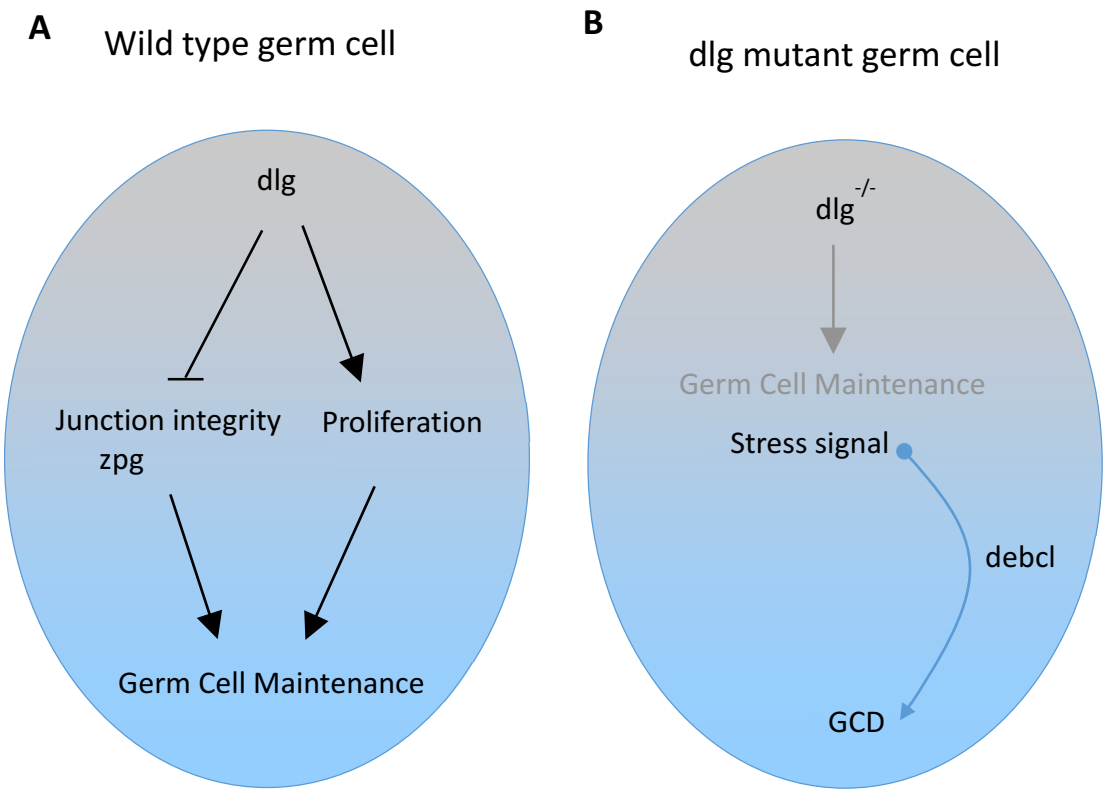



**Fig 5.11: dlg contributes to germ cell maintenance in different ways.** A) In wild type germ cells dlg promotes normal proliferation and maintains the junctional integrity required for proper cell–cell contact through its interaction E-cadherin and zpg. B) In the absence of dlg, germ cells behave aberrantly; proliferation is stunted and junctional integrity is compromised. As a result they release a stress signal which causes them to be eliminated by debcl-mediated cell death. GCD: germ cell death.

# Appendices

**Appendix A**: List of genes with significant differential expression in the Sxl-RNAi dataset. Adjusted p_value cutoff: <0.05

| gene id | gene name | Contr RNAi FPKM | Sxl-RNAi FPKM | log2(fold change) Sxl/contr | q value |
|---|---|---|---|---|---|
| FBgn0259923 | Sep4 | 25.9339 | 39.8976 | 0.621464 | 0.0282978 |
| FBgn0000427 | dec-1 | 51.3539 | 90.7103 | 0.820794 | 0.0040296 |
| FBgn0261504 | 7SLRNA:CR42652 | 111.318 | 218.856 | 0.975299 | 0.0131686 |
| FBgn0052865 | alphagamma-element:CR32865 | 288.959 | 37.8184 | -2.9337 | 0.0040296 |
| FBgn0003863 | alphaTry | 20.1949 | 304.896 | 3.91625 | 0.0040296 |
| FBgn0087040 | alphaTub67C | 4.31148 | 11.2966 | 1.38964 | 0.0040296 |
| FBgn0029879 | Apc7 | 1.85586 | 8.15311 | 2.13526 | 0.0040296 |
| FBgn0033926 | Arc1 | 77.1095 | 38.6364 | -0.996947 | 0.0040296 |
| FBgn0012042 | AttA | 97.9785 | 7.49021 | -3.70939 | 0.0040296 |
| FBgn0041581 | AttB | 47.4424 | 6.27149 | -2.9193 | 0.0040296 |
| FBgn0041579 | AttC | 29.379 | 6.67944 | -2.13699 | 0.0040296 |
| FBgn0032049 | Bace | 18.6445 | 221.275 | 3.56902 | 0.0040296 |
| FBgn0026149 | BCL7-like | 31.4282 | 65.1468 | 1.05164 | 0.0040296 |
| FBgn0010357 | betaTry | 21.8066 | 293.439 | 3.75023 | 0.0040296 |
| FBgn0000242 | Bx | 13.8966 | 26.4918 | 0.930815 | 0.0040296 |
| FBgn0260455 | CG10332 | 7.93251 | 1.52875 | -2.37543 | 0.0131686 |
| FBgn0038395 | CG10407 | 18.5742 | 55.0263 | 1.56682 | 0.0040296 |
| FBgn0039326 | CG10562 | 1.02366 | 3.65771 | 1.8372 | 0.0235649 |
| FBgn0035619 | CG10592 | 1.79254 | 9.68476 | 2.43371 | 0.0040296 |
| FBgn0046302 | CG10650 | 1.27899 | 5.41356 | 2.08158 | 0.0040296 |
| FBgn0032752 | CG10702 | 2.80298 | 6.90972 | 1.30167 | 0.0156429 |
| FBgn0029664 | CG10802 | 15.332 | 28.5982 | 0.899383 | 0.0131686 |
| FBgn0034295 | CG10911 | 3.31242 | 28.5207 | 3.10605 | 0.0040296 |
| FBgn0034296 | CG10912 | 0.504453 | 3.90699 | 2.95327 | 0.01035 |
| FBgn0031849 | CG11327 | 4.76797 | 16.2833 | 1.77195 | 0.0040296 |
| FBgn0035490 | CG1136 | 33.0082 | 55.7483 | 0.756102 | 0.0040296 |
| FBgn0039316 | CG11893 | 3.18931 | 12.6546 | 1.98834 | 0.0040296 |
| FBgn0037387 | CG1213 | 5.42441 | 12.052 | 1.15174 | 0.0156429 |

| | | | | | |
|---|---|---|---|---|---|
| FBgn0033774 | CG12374 | 10.1165 | 55.4173 | 2.45362 | 0.0040296 |
| FBgn0029825 | CG12728 | 22.4555 | 48.9964 | 1.12561 | 0.0040296 |
| FBgn0035694 | CG13299 | 657.215 | 1188.61 | 0.854836 | 0.0040296 |
| FBgn0033787 | CG13321 | 2.47558 | 10.2493 | 2.04968 | 0.0040296 |
| FBgn0033788 | CG13323 | 3.43424 | 14.7956 | 2.10711 | 0.00751791 |
| FBgn0033792 | CG13325 | 0.105603 | 1.76309 | 4.06138 | 0.0235649 |
| FBgn0034662 | CG13492 | 0.221631 | 1.2618 | 2.50925 | 0.0040296 |
| FBgn0035583 | CG13704 | 0.553849 | 3.15536 | 2.51024 | 0.0484327 |
| FBgn0034501 | CG13868 | 83.99 | 50.1216 | -0.744786 | 0.01035 |
| FBgn0035168 | CG13889 | 1.71578 | 0.396691 | -2.11277 | 0.0040296 |
| FBgn0035176 | CG13905 | 0.967122 | 5.07871 | 2.39269 | 0.0352048 |
| FBgn0031763 | CG13996 | 0.626095 | 0 | #NAME? | 0.0040296 |
| FBgn0030994 | CG14193 | 22.3515 | 49.4339 | 1.14513 | 0.0040296 |
| FBgn0038419 | CG14879 | 26.3828 | 41.0268 | 0.636969 | 0.0466844 |
| FBgn0030929 | CG15043 | 1.81368 | 9.35496 | 2.36681 | 0.0257574 |
| FBgn0034390 | CG15093 | 15.2835 | 33.6245 | 1.13754 | 0.0040296 |
| FBgn0028950 | CG15255 | 0.360805 | 3.59476 | 3.3166 | 0.01035 |
| FBgn0030667 | CG15599 | 2.10096 | 5.46642 | 1.37955 | 0.0040296 |
| FBgn0029766 | CG15784 | 68.0882 | 44.1276 | -0.625725 | 0.0411184 |
| FBgn0029754 | CG15930 | 10.3649 | 179.785 | 4.1165 | 0.0040296 |
| FBgn0033188 | CG1600 | 176.768 | 101.429 | -0.801391 | 0.0235649 |
| FBgn0030482 | CG1673 | 21.8752 | 38.8863 | 0.82997 | 0.0040296 |
| FBgn0031176 | CG1678 | 10.0874 | 1.3853 | -2.86428 | 0.0484327 |
| FBgn0032505 | CG16826 | 6.85749 | 0.536734 | -3.6754 | 0.0040296 |
| FBgn0042186 | CG17239 | 1.47647 | 7.81114 | 2.40338 | 0.0304434 |
| FBgn0034883 | CG17664 | 0.502052 | 2.92559 | 2.54282 | 0.0156429 |
| FBgn0035642 | CG18586 | 8.79239 | 18.1226 | 1.04346 | 0.0131686 |
| FBgn0030066 | CG1885 | 20.8923 | 40.6783 | 0.961287 | 0.0040296 |
| FBgn0033204 | CG2065 | 32.7253 | 15.7212 | -1.05769 | 0.0040296 |
| FBgn0023526 | CG2865 | 10.0517 | 20.3341 | 1.01647 | 0.0040296 |
| FBgn0050025 | CG30025 | 0.507964 | 4.14677 | 3.02919 | 0.0282978 |
| FBgn0050031 | CG30031 | 0.360846 | 4.33547 | 3.58673 | 0.0423906 |
| FBgn0050371 | CG30371 | 0.424508 | 2.61337 | 2.62205 | 0.0257574 |
| FBgn0051004 | CG31004 | 0.431821 | 1.89373 | 2.13273 | 0.0040296 |
| FBgn0051087 | CG31087 | 0.207461 | 2.35446 | 3.50448 | 0.0185411 |
| FBgn0051104 | CG31104 | 0.177085 | 1.79409 | 3.34074 | 0.01035 |
| FBgn0051198 | CG31198 | 2.46494 | 6.6484 | 1.43146 | 0.0040296 |
| FBgn0051233 | CG31233 | 1.81149 | 7.45746 | 2.04151 | 0.0040296 |
| FBgn0051262 | CG31262 | 1.68583 | 0.368284 | -2.19457 | 0.00751791 |
| FBgn0051681 | CG31681 | 1.1751 | 17.2549 | 3.87615 | 0.0040296 |
| FBgn0052204 | CG32204 | 2.30245 | 5.32594 | 1.20987 | 0.0156429 |

| | | | | | |
|---|---|---|---|---|---|
| FBgn0052333 | CG32333 | 1.34959 | 0.3858 | -1.8066 | 0.01035 |
| FBgn0052633 | CG32633 | 0.1694 | 2.81551 | 4.0549 | 0.00751791 |
| FBgn0052706 | CG32706 | 13.0726 | 30.4496 | 1.21988 | 0.00751791 |
| FBgn0052834 | CG32834 | 0.978816 | 11.9943 | 3.61517 | 0.0040296 |
| FBgn0053258 | CG33258 | 8.35195 | 0.459369 | -4.18439 | 0.0040296 |
| FBgn0250904 | CG34434 | 2.50872 | 11.8971 | 2.24558 | 0.0040296 |
| FBgn0036422 | CG3868 | 3.04325 | 37.9169 | 3.63916 | 0.0040296 |
| FBgn0040396 | CG3939 | 63.8588 | 98.3678 | 0.623302 | 0.0484327 |
| FBgn0038291 | CG3984 | 11.9431 | 2.23668 | -2.41675 | 0.0040296 |
| FBgn0263830 | CG40486 | 48.1113 | 23.1274 | -1.05678 | 0.0040296 |
| FBgn0029738 | CG4068 | 57.0098 | 37.4405 | -0.606609 | 0.0131686 |
| FBgn0259101 | CG42249 | 1.72437 | 5.20335 | 1.59337 | 0.0304434 |
| FBgn0259150 | CG42265 | 0.630417 | 1.78514 | 1.50166 | 0.0040296 |
| FBgn0261615 | CG42704 | 4.81829 | 194.643 | 5.33616 | 0.0040296 |
| FBgn0262009 | CG42827 | 0.666848 | 0 | #NAME? | 0.0040296 |
| FBgn0036787 | CG4306 | 6.19357 | 14.6567 | 1.24272 | 0.0423906 |
| FBgn0262536 | CG43090 | 5.34141 | 69.1267 | 3.69395 | 0.0040296 |
| FBgn0262719 | CG43163 | 1.74238 | 0.467092 | -1.89928 | 0.0040296 |
| FBgn0034663 | CG4363 | 0.589418 | 15.7102 | 4.73627 | 0.0040296 |
| FBgn0034664 | CG4377 | 0.682725 | 8.57084 | 3.65006 | 0.0040296 |
| FBgn0038358 | CG4525 | 0 | 0.644413 | inf | 0.0040296 |
| FBgn0040723 | CG5011 | 1.15322 | 12.2865 | 3.41334 | 0.0131686 |
| FBgn0039342 | CG5107 | 5.86432 | 130.731 | 4.4785 | 0.0040296 |
| FBgn0035620 | CG5150 | 1.56785 | 7.29821 | 2.21876 | 0.0040296 |
| FBgn0038082 | CG5724 | 0.18171 | 4.39704 | 4.59682 | 0.0040296 |
| FBgn0032587 | CG5953 | 14.0293 | 8.39768 | -0.74038 | 0.0393516 |
| FBgn0038349 | CG6045 | 33.1068 | 19.3367 | -0.775787 | 0.0040296 |
| FBgn0031918 | CG6055 | 61.0785 | 29.8813 | -1.03142 | 0.0040296 |
| FBgn0030647 | CG6324 | 0.598787 | 1.7696 | 1.56331 | 0.0444441 |
| FBgn0028740 | CG6362 | 2.98886 | 0.947301 | -1.6577 | 0.0330295 |
| FBgn0029693 | CG6379 | 15.5702 | 29.3903 | 0.916549 | 0.0040296 |
| FBgn0034247 | CG6484 | 5.22345 | 12.5041 | 1.25932 | 0.0156429 |
| FBgn0035923 | CG6511 | 18.5353 | 11.5262 | -0.685358 | 0.0212404 |
| FBgn0038290 | CG6912 | 24.0315 | 5.25355 | -2.19356 | 0.0040296 |
| FBgn0031971 | CG7224 | 185.568 | 94.3784 | -0.975423 | 0.0040296 |
| FBgn0036738 | CG7542 | 4.45441 | 65.5006 | 3.8782 | 0.0040296 |
| FBgn0034105 | CG7755 | 1.37514 | 0.17607 | -2.96535 | 0.0235649 |
| FBgn0033047 | CG7882 | 1.27551 | 0.17584 | -2.85873 | 0.0411184 |
| FBgn0038135 | CG8773 | 0.323222 | 2.37298 | 2.8761 | 0.00751791 |
| FBgn0033733 | CG8834 | 0.970852 | 9.8332 | 3.34034 | 0.0040296 |
| FBgn0034493 | CG8908 | 10.9073 | 4.624 | -1.23808 | 0.0040296 |

| | | | | | |
|---|---|---|---|---|---|
| FBgn0032066 | CG9463 | 0.183724 | 1.38234 | 2.9115 | 0.0040296 |
| FBgn0034807 | CG9897 | 1.80327 | 16.2364 | 3.17054 | 0.0040296 |
| FBgn0034829 | CG9899 | 2.97017 | 0.836225 | -1.82858 | 0.0257574 |
| FBgn0004859 | ci | 25.9286 | 40.5614 | 0.645563 | 0.0156429 |
| FBgn0030529 | Clic | 106.533 | 158.067 | 0.569238 | 0.0304434 |
| FBgn0263492 | CR43481 | 7.15783 | 1.87177 | -1.93512 | 0.0131686 |
| FBgn0015714 | Cyp6a17 | 25.0711 | 10.5974 | -1.24232 | 0.0040296 |
| FBgn0013771 | Cyp6a9 | 33.5709 | 68.4075 | 1.02694 | 0.0040296 |
| FBgn0015039 | Cyp9b2 | 4.41347 | 12.0274 | 1.44634 | 0.0040296 |
| FBgn0025641 | DAAM | 16.2166 | 25.0017 | 0.62455 | 0.0235649 |
| FBgn0011761 | dhd | 19.8858 | 59.4516 | 1.57998 | 0.0040296 |
| FBgn0085358 | Diedel3 | 4.20963 | 54.7172 | 3.70023 | 0.0040296 |
| FBgn0035331 | DmsR-1 | 5.69345 | 1.92463 | -1.56472 | 0.0040296 |
| FBgn0263106 | DnaJ-1 | 504.159 | 221.562 | -1.18617 | 0.00751791 |
| FBgn0037141 | DNApol-eta | 14.7624 | 23.9846 | 0.700181 | 0.0411184 |
| FBgn0260866 | dnr1 | 29.8682 | 17.6162 | -0.761706 | 0.0040296 |
| FBgn0004240 | Dpt | 27.9924 | 6.34116 | -2.14222 | 0.0040296 |
| FBgn0034407 | DptB | 81.8398 | 29.9634 | -1.4496 | 0.0040296 |
| FBgn0015380 | drl | 28.4608 | 45.2065 | 0.667558 | 0.00751791 |
| FBgn0010388 | Dro | 189.303 | 35.6685 | -2.40798 | 0.0040296 |
| FBgn0261046 | Dscam3 | 1.13858 | 0.0564918 | -4.33306 | 0.0040296 |
| FBgn0011764 | Dsp1 | 89.9581 | 149.512 | 0.732939 | 0.0040296 |
| FBgn0000535 | eag | 0.944377 | 0.319398 | -1.56401 | 0.0352048 |
| FBgn0004554 | Edg91 | 5.67132 | 1.08828 | -2.38164 | 0.0304434 |
| FBgn0035860 | eIF4E-3 | 71.046 | 133.345 | 0.908335 | 0.0040296 |
| FBgn0010425 | epsilonTry | 7.87525 | 91.1684 | 3.53314 | 0.0040296 |
| FBgn0033153 | Gadd45 | 60.2121 | 29.0196 | -1.05302 | 0.0040296 |
| FBgn0010359 | gammaTry | 3.75635 | 24.9037 | 2.72895 | 0.0040296 |
| FBgn0010038 | GstD2 | 12.3064 | 1.77521 | -2.79334 | 0.0040296 |
| FBgn0053801 | His1:CG33801 | 26.0535 | 52.5863 | 1.01321 | 0.01035 |
| FBgn0053803 | His3:CG33803 | 156.628 | 313.008 | 0.998856 | 0.0040296 |
| FBgn0051754 | His-Psi:CR31754 | 250.477 | 486.311 | 0.957201 | 0.0040296 |
| FBgn0001217 | Hsc70-2 | 23.5485 | 9.9184 | -1.24745 | 0.0040296 |
| FBgn0001223 | Hsp22 | 971.186 | 173.655 | -2.48353 | 0.0040296 |
| FBgn0001224 | Hsp23 | 2862.89 | 484.851 | -2.56186 | 0.0040296 |
| FBgn0001225 | Hsp26 | 3390.69 | 680.266 | -2.31741 | 0.0040296 |
| FBgn0001226 | Hsp27 | 1742.73 | 690.61 | -1.3354 | 0.01035 |
| FBgn0001229 | Hsp67Bc | 193.161 | 18.5054 | -3.38379 | 0.0040296 |
| FBgn0001230 | Hsp68 | 871.624 | 239.357 | -1.86454 | 0.0040296 |
| FBgn0013278 | Hsp70Bb | 30.4094 | 11.9443 | -1.34819 | 0.0040296 |

| | | | | | |
|---|---|---|---|---|---|
| FBgn0051354 | Hsp70Bbb | 32.7143 | 15.6201 | -1.06652 | 0.0040296 |
| FBgn0013279 | Hsp70Bc | 29.2442 | 11.7722 | -1.31276 | 0.0040296 |
| FBgn0001234 | Hsromega | 141.017 | 58.723 | -1.26387 | 0.0040296 |
| FBgn0010389 | htl | 29.5494 | 43.4901 | 0.55756 | 0.0423906 |
| FBgn0020906 | Jon25Bi | 3.94419 | 20.4824 | 2.37658 | 0.0040296 |
| FBgn0031654 | Jon25Bii | 1.4427 | 7.80879 | 2.43633 | 0.0040296 |
| FBgn0031653 | Jon25Biii | 2.22421 | 10.7263 | 2.26979 | 0.0040296 |
| FBgn0035665 | Jon65Aiii | 7.79164 | 60.3169 | 2.95256 | 0.0040296 |
| FBgn0250815 | Jon65Aiv | 18.3629 | 112.142 | 2.61046 | 0.0040296 |
| FBgn0039777 | Jon99Fii | 0.706509 | 3.84902 | 2.44571 | 0.0257574 |
| FBgn0011296 | l(2)efl | 237.09 | 134.293 | -0.820052 | 0.0040296 |
| FBgn0002565 | Lsp2 | 18.3368 | 9.24005 | -0.988773 | 0.0040296 |
| FBgn0002570 | Mal-A1 | 2.42645 | 58.955 | 4.60269 | 0.01035 |
| FBgn0002569 | Mal-A2 | 0.390408 | 2.56322 | 2.7149 | 0.0040296 |
| FBgn0050360 | Mal-A6 | 2.33732 | 18.7504 | 3.004 | 0.0040296 |
| FBgn0033296 | Mal-A7 | 1.05253 | 48.8398 | 5.53613 | 0.0040296 |
| FBgn0033297 | Mal-A8 | 0.777066 | 23.0294 | 4.88929 | 0.0040296 |
| FBgn0019985 | mGluRA | 7.74856 | 3.35661 | -1.20692 | 0.00751791 |
| FBgn0263112 | Mitf | 23.1362 | 35.8342 | 0.631182 | 0.0212404 |
| FBgn0014340 | mof | 21.3666 | 37.9352 | 0.828178 | 0.0040296 |
| FBgn0263316 | Mrp4 | 22.8823 | 13.6771 | -0.742468 | 0.0040296 |
| FBgn0038790 | MtnC | 2.89453 | 18.0589 | 2.64131 | 0.0393516 |
| FBgn0086347 | Myo31DF | 45.5746 | 30.7495 | -0.567666 | 0.0423906 |
| FBgn0043535 | Obp57a | 53.4961 | 130.91 | 1.29107 | 0.0040296 |
| FBgn0003015 | osk | 1.33931 | 0.373828 | -1.84104 | 0.0444441 |
| FBgn0035317 | osm-1 | 5.59682 | 2.49257 | -1.16697 | 0.0040296 |
| FBgn0038973 | Pebp1 | 8.00957 | 131.921 | 4.0418 | 0.0040296 |
| FBgn0263234 | Phae1 | 1.36616 | 4.95602 | 1.85906 | 0.0373111 |
| FBgn0263235 | Phae2 | 0.885445 | 5.76571 | 2.70302 | 0.0040296 |
| FBgn0004959 | phm | 18.0204 | 29.6165 | 0.716769 | 0.0156429 |
| FBgn0034647 | pirk | 8.52336 | 3.20551 | -1.41087 | 0.0393516 |
| FBgn0003231 | ref(2)P | 252.179 | 156.149 | -0.691519 | 0.01035 |
| FBgn0014018 | Rel | 106.64 | 52.0966 | -1.03348 | 0.0040296 |
| FBgn0034837 | RpL22-like | 31.5046 | 55.6661 | 0.821238 | 0.00751791 |
| FBgn0002306 | sas | 2.58305 | 0.92898 | -1.47536 | 0.0156429 |
| FBgn0033033 | scaf | 9.20658 | 2.71223 | -1.76319 | 0.0040296 |
| FBgn0031406 | Send1 | 1.43602 | 13.1785 | 3.19805 | 0.0040296 |
| FBgn0035539 | slow | 12.0444 | 19.9133 | 0.725368 | 0.0352048 |
| FBgn0039873 | Smvt | 1.6528 | 0.214848 | -2.94353 | 0.01035 |
| FBgn0065046 | snoRNA:U3:9B | 23.2069 | 116.422 | 2.32673 | 0.0373111 |
| FBgn0083987 | snRNA:U11 | 32.5601 | 151.637 | 2.21944 | 0.0423906 |

| | | | | | |
|---|---|---|---|---|---|
| FBgn0004191 | snRNA:U2:34ABa | 429.309 | 1022.73 | 1.25234 | 0.0040296 |
| FBgn0031973 | Spn28D | 26.3256 | 12.998 | -1.01817 | 0.0040296 |
| FBgn0029768 | SPR | 9.54465 | 17.1417 | 0.844747 | 0.01035 |
| FBgn0051641 | stai | 234.684 | 358.694 | 0.612038 | 0.0373111 |
| FBgn0024836 | stan | 1.64416 | 3.3407 | 1.0228 | 0.0040296 |
| FBgn0086708 | stv | 209.635 | 76.392 | -1.45639 | 0.0040296 |
| FBgn0264270 | Sxl | 176.294 | 63.124 | -1.48172 | 0.0040296 |
| FBgn0031623 | Taf12L | 1.73779 | 7.55338 | 2.11986 | 0.0257574 |
| FBgn0010329 | Tbh | 2.00427 | 6.18669 | 1.62609 | 0.0040296 |
| FBgn0041182 | TepII | 16.0946 | 9.20985 | -0.805331 | 0.0040296 |
| FBgn0011555 | thetaTry | 2.13958 | 14.0146 | 2.71153 | 0.0040296 |
| FBgn0261575 | tobi | 21.318 | 45.6759 | 1.09936 | 0.0040296 |
| FBgn0032074 | Tsp29Fa | 1.50032 | 5.82876 | 1.95792 | 0.0212404 |
| FBgn0031424 | VGlut | 0.358184 | 1.61667 | 2.17426 | 0.0411184 |
| FBgn0032373 | Vha100-5 | 0.373887 | 2.26347 | 2.59786 | 0.0185411 |
| FBgn0039896 | yellow-h | 10.8534 | 23.4444 | 1.1111 | 0.01035 |
| FBgn0015565 | yin | 11.3293 | 6.32187 | -0.841641 | 0.0423906 |

**Appendix B:** List of isoforms with significant differential expression in the Sxl-RNAi dataset. Adjusted p_value cutoff: <0.05

| Transcript id | gene name | Contr RNAi FPKM | Sxl-RNAi FPKM | log2(fold change) Sxl/contr | q_value |
|---|---|---|---|---|---|
| FBtr0302398 | 7SLRNA:CR42652 | 111.318 | 218.856 | 0.975299 | 0.025758 |
| FBtr0082678 | alphagamma-element:CR32865 | 288.959 | 37.8184 | -2.9337 | 0.007621 |
| FBtr0088161 | alphaTry | 20.1949 | 304.896 | 3.91625 | 0.007621 |
| FBtr0076393 | alphaTub67C | 4.31148 | 11.2966 | 1.38964 | 0.007621 |
| FBtr0087560 | Arc1 | 77.1095 | 38.6364 | -0.996947 | 0.007621 |
| FBtr0087437 | AttA | 97.9785 | 7.49021 | -3.70939 | 0.007621 |
| FBtr0087438 | AttB | 43.1219 | 4.81163 | -3.16382 | 0.007621 |
| FBtr0305795 | AttC | 29.379 | 6.67944 | -2.13699 | 0.007621 |
| FBtr0079701 | Bace | 18.6445 | 221.275 | 3.56902 | 0.007621 |
| FBtr0071358 | BCL7-like | 31.4282 | 65.1468 | 1.05164 | 0.007621 |
| FBtr0088122 | betaTry | 21.8066 | 293.439 | 3.75023 | 0.007621 |
| FBtr0072109 | CG10332 | 7.93251 | 1.52875 | -2.37543 | 0.025758 |
| FBtr0083237 | CG10407 | 18.5742 | 55.0263 | 1.56682 | 0.007621 |
| FBtr0084851 | CG10562 | 1.02366 | 3.65771 | 1.8372 | 0.047132 |
| FBtr0077129 | CG10592 | 1.79254 | 9.68476 | 2.43371 | 0.007621 |
| FBtr0081183 | CG10650 | 1.27899 | 5.41356 | 2.08158 | 0.007621 |
| FBtr0086808 | CG10911 | 3.31242 | 28.5207 | 3.10605 | 0.007621 |
| FBtr0086807 | CG10912 | 0.504453 | 3.90699 | 2.95327 | 0.019779 |
| FBtr0079315 | CG11327 | 4.76797 | 16.2833 | 1.77195 | 0.007621 |
| FBtr0073203 | CG1136 | 33.0082 | 55.7483 | 0.756102 | 0.007621 |
| FBtr0084827 | CG11893 | 3.18931 | 12.6546 | 1.98834 | 0.007621 |
| FBtr0087854 | CG12374 | 10.1165 | 55.4173 | 2.45362 | 0.007621 |
| FBtr0070859 | CG12728 | 22.4555 | 48.9964 | 1.12561 | 0.007621 |
| FBtr0077013 | CG13299 | 657.215 | 1188.61 | 0.854836 | 0.007621 |
| FBtr0087792 | CG13321 | 2.47558 | 10.2493 | 2.04968 | 0.007621 |
| FBtr0087797 | CG13323 | 3.43424 | 14.7956 | 2.10711 | 0.014574 |
| FBtr0087722 | CG13325 | 0.105603 | 1.76309 | 4.06138 | 0.047132 |
| FBtr0302446 | CG13492 | 0.221631 | 1.2618 | 2.50925 | 0.007621 |
| FBtr0301209 | CG13868 | 83.6611 | 50.0898 | -0.740039 | 0.019779 |

| | | | | | |
|---|---|---|---|---|---|
| FBtr0079181 | CG13996 | 0.626095 | 0 | #NAME? | 0.007621 |
| FBtr0074670 | CG14193 | 22.3515 | 49.4339 | 1.14513 | 0.007621 |
| FBtr0080785 | CG15255 | 0.360805 | 3.59476 | 3.3166 | 0.019779 |
| FBtr0307503 | CG15599 | 2.10096 | 5.46642 | 1.37955 | 0.007621 |
| FBtr0301881 | CG15930 | 10.3649 | 179.785 | 4.1165 | 0.007621 |
| FBtr0073777 | CG1673 | 21.8752 | 38.8863 | 0.82997 | 0.007621 |
| FBtr0080464 | CG16826 | 6.85749 | 0.536734 | -3.6754 | 0.007621 |
| FBtr0072107 | CG17664 | 0.502052 | 2.92559 | 2.54282 | 0.031229 |
| FBtr0300903 | CG18586 | 8.40473 | 17.4523 | 1.05414 | 0.019779 |
| FBtr0088922 | CG2065 | 32.7253 | 15.7212 | -1.05769 | 0.007621 |
| FBtr0084855 | CG31087 | 0.207461 | 2.35446 | 3.50448 | 0.03692 |
| FBtr0084825 | CG31104 | 0.177085 | 1.79409 | 3.34074 | 0.019779 |
| FBtr0084128 | CG31198 | 2.46494 | 6.6484 | 1.43146 | 0.007621 |
| FBtr0084127 | CG31233 | 1.81149 | 7.45746 | 2.04151 | 0.007621 |
| FBtr0077802 | CG31681 | 1.1751 | 17.2549 | 3.87615 | 0.007621 |
| FBtr0113429 | CG32204 | 2.30245 | 5.32594 | 1.20987 | 0.031229 |
| FBtr0071287 | CG32706 | 13.0726 | 30.4496 | 1.21988 | 0.014574 |
| FBtr0303210 | CG32834 | 0.978816 | 11.9943 | 3.61517 | 0.007621 |
| FBtr0301030 | CG33258 | 8.35195 | 0.459369 | -4.18439 | 0.007621 |
| FBtr0112735 | CG34434 | 2.50872 | 11.8971 | 2.24558 | 0.007621 |
| FBtr0075729 | CG3868 | 3.04325 | 37.9169 | 3.63916 | 0.007621 |
| FBtr0083025 | CG3984 | 11.9431 | 2.23668 | -2.41675 | 0.007621 |
| FBtr0111188 | CG40486 | 47.6044 | 22.3551 | -1.09049 | 0.007621 |
| FBtr0302954 | CG42704 | 4.81829 | 194.643 | 5.33616 | 0.007621 |
| FBtr0304876 | CG43090 | 5.34141 | 69.1267 | 3.69395 | 0.007621 |
| FBtr0071740 | CG4363 | 0.589418 | 15.7102 | 4.73627 | 0.007621 |
| FBtr0071739 | CG4377 | 0.682725 | 8.57084 | 3.65006 | 0.007621 |
| FBtr0083170 | CG4525 | 0 | 0.644413 | inf | 0.007621 |
| FBtr0077973 | CG5011 | 1.15322 | 12.2865 | 3.41334 | 0.025758 |
| FBtr0084879 | CG5107 | 5.86432 | 130.731 | 4.4785 | 0.007621 |
| FBtr0077086 | CG5150 | 1.56785 | 7.29821 | 2.21876 | 0.007621 |
| FBtr0082657 | CG5724 | 0.18171 | 4.39704 | 4.59682 | 0.007621 |
| FBtr0083151 | CG6045 | 33.1068 | 19.3367 | -0.775787 | 0.007621 |
| FBtr0079449 | CG6055 | 60.9774 | 29.8813 | -1.02903 | 0.007621 |
| FBtr0070671 | CG6379 | 15.5702 | 29.3903 | 0.916549 | 0.007621 |
| FBtr0086903 | CG6484 | 5.22345 | 12.5041 | 1.25932 | 0.031229 |
| FBtr0083038 | CG6912 | 24.0315 | 5.25355 | -2.19356 | 0.007621 |
| FBtr0079532 | CG7224 | 182.182 | 92.0358 | -0.985112 | 0.007621 |
| FBtr0075181 | CG7542 | 4.45441 | 65.5006 | 3.8782 | 0.007621 |
| FBtr0087182 | CG7755 | 1.37514 | 0.17607 | -2.96535 | 0.047132 |
| FBtr0290037 | CG8773 | 0.323222 | 2.37298 | 2.8761 | 0.014574 |

| | | | | | |
|---|---|---|---|---|---|
| FBtr0087954 | CG8834 | 0.942269 | 9.80028 | 3.37861 | 0.007621 |
| FBtr0332256 | CG8908 | 10.9073 | 4.624 | -1.23808 | 0.007621 |
| FBtr0079744 | CG9463 | 0.183724 | 1.38234 | 2.9115 | 0.007621 |
| FBtr0333547 | CG9576 | 4.3371 | 16.8739 | 1.95999 | 0.007621 |
| FBtr0071989 | CG9897 | 1.80327 | 16.2364 | 3.17054 | 0.007621 |
| FBtr0089178 | ci | 25.216 | 39.6673 | 0.653614 | 0.042445 |
| FBtr0309503 | CR43481 | 5.84165 | 1.75246 | -1.73699 | 0.047132 |
| FBtr0087451 | Cyp6a17 | 25.0711 | 10.5974 | -1.24232 | 0.007621 |
| FBtr0300128 | Cyp6a9 | 33.5709 | 68.4075 | 1.02694 | 0.007621 |
| FBtr0089055 | Cyp9b2 | 4.41347 | 12.0274 | 1.44634 | 0.007621 |
| FBtr0070239 | DAAM | 11.2338 | 19.6143 | 0.80406 | 0.014574 |
| FBtr0070749 | dhd | 19.8858 | 59.4516 | 1.57998 | 0.007621 |
| FBtr0112531 | Diedel3 | 4.20963 | 54.7172 | 3.70023 | 0.007621 |
| FBtr0330187 | DmsR-1 | 5.69345 | 1.92463 | -1.56472 | 0.007621 |
| FBtr0077123 | DnaJ-1 | 486.645 | 214.242 | -1.18363 | 0.019779 |
| FBtr0330229 | dnr1 | 29.8682 | 17.6161 | -0.761716 | 0.007621 |
| FBtr0086620 | Dpt | 27.9924 | 6.34116 | -2.14222 | 0.007621 |
| FBtr0086621 | DptB | 81.8398 | 29.9634 | -1.4496 | 0.007621 |
| FBtr0081195 | drl | 28.4608 | 45.2065 | 0.667558 | 0.014574 |
| FBtr0330211 | Dscam3 | 0.692677 | 0.048881 | -3.82483 | 0.007621 |
| FBtr0089262 | Dsp1 | 51.7836 | 92.0676 | 0.830198 | 0.042445 |
| FBtr0076701 | eIF4E-3 | 71.046 | 133.345 | 0.908335 | 0.007621 |
| FBtr0088160 | epsilonTry | 7.87525 | 91.1684 | 3.53314 | 0.007621 |
| FBtr0089049 | Gadd45 | 60.2121 | 29.0196 | -1.05302 | 0.007621 |
| FBtr0088159 | gammaTry | 3.75635 | 24.9037 | 2.72895 | 0.007621 |
| FBtr0082569 | GstD2 | 12.3064 | 1.77521 | -2.79334 | 0.007621 |
| FBtr0091805 | His1:CG33801 | 26.0535 | 52.5863 | 1.01321 | 0.019779 |
| FBtr0091807 | His3:CG33803 | 156.628 | 313.008 | 0.998856 | 0.007621 |
| FBtr0085891 | His-Psi:CR31754 | 250.477 | 486.311 | 0.957201 | 0.007621 |
| FBtr0082707 | Hsc70-2 | 23.5485 | 9.9184 | -1.24745 | 0.007621 |
| FBtr0309504 | Hsp23 | 1773.06 | 300.446 | -2.56106 | 0.007621 |
| FBtr0076496 | Hsp26 | 3390.69 | 680.266 | -2.31741 | 0.007621 |
| FBtr0076454 | Hsp27 | 1742.73 | 690.61 | -1.3354 | 0.019779 |
| FBtr0076497 | Hsp67Bc | 193.161 | 18.5054 | -3.38379 | 0.007621 |
| FBtr0084589 | Hsp68 | 871.624 | 239.357 | -1.86454 | 0.007621 |
| FBtr0082637 | Hsp70Bb | 30.4094 | 11.9443 | -1.34819 | 0.007621 |
| FBtr0082636 | Hsp70Bbb | 32.7143 | 15.6201 | -1.06652 | 0.007621 |

| | | | | | |
|---|---|---|---|---|---|
| FBtr0082638 | Hsp70Bc | 29.2442 | 11.7722 | -1.31276 | 0.007621 |
| FBtr0100432 | Jon25Bi | 3.94419 | 20.4824 | 2.37658 | 0.007621 |
| FBtr0079055 | Jon25Bii | 1.4427 | 7.80879 | 2.43633 | 0.007621 |
| FBtr0079056 | Jon25Biii | 2.22421 | 10.7263 | 2.26979 | 0.007621 |
| FBtr0077041 | Jon65Aiii | 7.79164 | 60.3169 | 2.95256 | 0.007621 |
| FBtr0077040 | Jon65Aiv | 18.3629 | 112.142 | 2.61046 | 0.007621 |
| FBtr0072100 | l(2)efl | 235.862 | 133.936 | -0.816394 | 0.007621 |
| FBtr0089324 | Lsp2 | 18.3368 | 9.24005 | -0.988773 | 0.007621 |
| FBtr0088759 | Mal-A1 | 2.42645 | 58.955 | 4.60269 | 0.019779 |
| FBtr0088747 | Mal-A2 | 0.390408 | 2.56322 | 2.7149 | 0.007621 |
| FBtr0273322 | Mal-A6 | 2.33732 | 18.7504 | 3.004 | 0.007621 |
| FBtr0088757 | Mal-A7 | 1.05253 | 48.8398 | 5.53613 | 0.007621 |
| FBtr0088752 | Mal-A8 | 0.777066 | 23.0294 | 4.88929 | 0.007621 |
| FBtr0070829 | mof | 21.3666 | 37.9352 | 0.828178 | 0.007621 |
| FBtr0302027 | Mrp4 | 22.8823 | 13.6771 | -0.742468 | 0.007621 |
| FBtr0086294 | Obp57a | 53.4961 | 130.91 | 1.29107 | 0.007621 |
| FBtr0332661 | osm-1 | 5.59682 | 2.49257 | -1.16697 | 0.007621 |
| FBtr0084255 | Pebp1 | 8.00957 | 131.921 | 4.0418 | 0.007621 |
| FBtr0302854 | Phae2 | 0.885445 | 5.76571 | 2.70302 | 0.007621 |
| FBtr0074603 | phm | 18.0204 | 29.6165 | 0.716769 | 0.031229 |
| FBtr0072049 | RpL22-like | 30.4139 | 53.9701 | 0.827428 | 0.007621 |
| FBtr0086060 | scaf | 8.72357 | 2.61257 | -1.73945 | 0.007621 |
| FBtr0077804 | Send1 | 1.43602 | 13.1785 | 3.19805 | 0.007621 |
| FBtr0085860 | Smvt | 1.6528 | 0.214848 | -2.94353 | 0.019779 |
| FBtr0080486 | snRNA:U2:34ABa | 429.309 | 1022.73 | 1.25234 | 0.007621 |
| FBtr0079549 | Spn28D | 26.3256 | 12.998 | -1.01817 | 0.007621 |
| FBtr0070778 | SPR | 9.54465 | 17.1417 | 0.844747 | 0.019779 |
| FBtr0303223 | stan | 1.12372 | 3.10835 | 1.46786 | 0.007621 |
| FBtr0331264 | Sxl | 82.3407 | 21.6879 | -1.92471 | 0.007621 |
| FBtr0089999 | Tbh | 2.00427 | 6.18669 | 1.62609 | 0.007621 |
| FBtr0088162 | thetaTry | 2.13958 | 14.0146 | 2.71153 | 0.007621 |
| FBtr0084847 | tobi | 21.318 | 45.6759 | 1.09936 | 0.007621 |
| FBtr0080255 | Vha100-5 | 0.373887 | 2.26347 | 2.59786 | 0.03692 |
| FBtr0089115 | yellow-h | 10.8534 | 23.4444 | 1.1111 | 0.019779 |

**Appendix C:** List of primers used for sex-specific alternative exon usage screen. Highlighted rows represent control primers used to compare the expression of proposed sex-specific exons to common (control) exons.

| Gene Symbol | Exon boundaries | Left Primer | Right Primer |
|---|---|---|---|
| 14-3-3zeta | CG17870.5 - CG17870.13 | CCTTATTCCAAAAGCCAGCA | CCCTCGCTGGTGAATTGATA |
| 14-3-3zeta | CG17870.4 - CG17870.5 | CGAAGCATCCGCTAGAAAAC | TGGCAACCTCGGCTAAATAC |
| Asph | CG8421.8 - CG8421.9 | ATACCGGATGATGTGGAGGA | ATCGCTAACTCCGTCATGCT |
| Asph | CG8421.3 - CG8421.4 | ATCACGATGAACACGACGAC | GCTTCCGTAGTGGCTTCAGT |
| Atpalpha | CG5670.6 - CG5670.17 | TGGTGTGGAGGAGGGTCGCC | GGGTTCCTCGGCTGGCGTTTC |
| Atpalpha | CG5670.17 - CG5670.8-A/B | AAACGCCAGCCGAGGAACCC | TGGGCAAGAAGCCGTTCTCAGC |
| Atpalpha | CG5670.3 - CG5670.4 | ACGCAAGATGCCGGCCAAAG | GCATTGGGACCATCGCGCTC |
| betaTub97EF | CG4869.5 - CG4869.4 | CTTGGCTACTCACTGACACC | TTGACAACGAGGCTCTCTAC |
| betaTub97EF | CG4869.3 - CG4869.2 | GGAGAAGGAGTTCATGATGC | AGATAACTTCGTTTACGGACAG |
| bun | FBgn0010460.4 - FBgn0010460.5 | GTTATCAATCGCAACTGCAC | ATTAAACAAGCGTGAAAAACC |
| bun | FBgn0010460.2 - FBgn0010460.4 | GGATGTTCGACTTGAGAATG | GTGCAGTTGCGATTGATAAC |
| ced-6 | CG11804.2 - CG11804.9 | GATAAGGCATTGTTGTGGTC | GCTGCTCCTACTCCTACTCC |
| ced-6 | CG11804.3 - CG11804.2 | TGCTAATGGTGATCTCAAGC | TTCTGGAACAAGCAGAACAG |
| CG11961 | CG11961.1 - CG11961.13 | GGCAGTGAGAAGAATGTACGG | CAGCCAGAATCCAAAGAAGG |
| CG11961 | CG11961.13 - CG11961.3 | GCCACTAAAGAACCCTGTGG | CATTAGCCAGGGATGATTGG |
| CG12065 | CG12065.5 - CG12065.6 | CTACTGCTACGGCGACTGC | GAGCTCCAGCACCTTCTCC |
| CG12065 | CG12065.4 - CG12065.5 | CTCCTGCAGCACGGATACTC | CGCCGTAGCAGTAGATGATG |
| CG12065 | CG12065.6 - CG12065.7 | AGCAGGCATATCCCAATCC | GATCACCAGCAGCACATCC |

| | | | |
|---|---|---|---|
| CG17370 | CG17370.9 - CG17370.2 | AATTTCTCAAGCGCTGCAAG | TTCTTCGCTTATCGGTCGTC |
| CG17370 | CG17370.3 - CG17370.4 | CCTCGGAGCTTCCATCTCAC | ATGGCCTGATGATGTACTGG |
| CG31522 | CG31522.272254-272147 - CG31522.271968-271888 | TTCAGAGGCTTTCGGTTCTC | GCCGGTGGACTACACAGATAG |
| CG31522 | CG31522.269779-269624 - CG31522.268777-267137 | TTGCAATCGATGAAGAGCAG | TTAGGTGGAAATGGAGTGGAG |
| CG32560 | FBgn0052560.1 - FBgn0052560.3 | CGCGCGAGAGAAAAGCCGAG | CGGCGGTCCGTCCCACAAAG |
| CG32560 | FBgn0052560.7 - FBgn0052560.8 | TGCCGCGGGACACCATCAAC | GCGAAGCTCGGCCTCATCCC |
| CG33275 | FBgn0035802.4 - FBgn0035802.5-B | TTCCTCGCCGAATAGTTTTG | CAGTGCTCCATTGGATGTTG |
| CG33275 | FBgn0035802.6-A - FBgn0035802.7 | AGCGTTGGTAATGAGGCAAG | ATTGAGCTGGACGGCAGTAG |
| CG31919 | CG31919.19 – CG31919.2 | TACGGGGGAATCTAGCCTCC | CGCCGATAAAATCCGACAAGG |
| CG31919 | CG31919.5 - CG31919.6 | CGGCCCGCAATGGGAACGAG | GTGGCCCCGCATCACTGCTC |
| CG34394 | FBgn0085423.3 - FBgn0085423.4 | GTGATGATGATGCTGACTACC | AGGGCTTCATTGAGAAGAAC |
| CG34394 | FBgn0085423.8 - FBgn0085423.9 | GTTGTAGGGCACACGAGTAG | AGTCCTGGTACAACGTCATC |
| CG34417 | FBgn0085446.8 - FBgn0085446.9 | AAGATCACTCGCACAATGAC | CTGCTCCTCGATCACATAAC |
| CG34417 | FBgn0085446.9 - FBgn0085446.11 | TGCGTGAGTTCAAGAAGGTG | TGTTATTGGTGGTGCTGGTG |
| CG34417 | FBgn0085446.12 - FBgn0085446.13 | CACGTCTACGCGAACTTATGG | CTCCTGTGGCTCCTGTTTTC |
| CG34417 | FBgn0085446.11 - FBgn0085446.12 | TTGAGGAGATTTGGGACGAG | CGCCATAAGTTCGCGTAGAC |
| CG3558 | CG3558.5 - CG3558.6 | CAGCAAATCCATTACCAGCA | GCGTAAATTTGTGATAAGCCAGA |
| CG3558 | CG3558.4 - CG3558.5 | TTGCGCAGGGCACCGTAAGC | AGCGACGAACGCGCCACATC |
| CG5850 | FBgn0032172.3 - FBgn0032172.4 | CGGCAAGTGTAGTTGTTGTT | CAGCAGTGAGTCCGAGTATC |
| CG5973 | CG5973.9 - CG5973.2 | ATTTCCGCTTGCCAGTGAC | TCGATGTAGGGCAACTTGTG |

| | | | |
|---|---|---|---|
| CG5973 | CG5973.1 - CG5973.2 | TCGAGGCGTGTAATTTTGTG | TCGATGTAGGGCAACTTGTG |
| CG5973 | CG5973.2 - CG5973.3 | AGGCCATCAAGGAACTCAGG | CATCCAATGGCAGGTTCAG |
| CG6043 | FBgn0032497.1 - FBgn0032497.3 | GACTTGGGCTGCTAACTTG | CTTTTACACACCGCCTACAC |
| CG6043 | FBgn0032497.2 - FBgn0032497.5 | GTAGGCGGTGTGTAAAAGTG | AGGGCTTCTTGTCTTTTGTC |
| CG6043 | FBgn0032497.14 - FBgn0032497.15 | ATGACAACACCTGAAAATGC | TCGCGTTCAACAATAAAAAC |
| CG6043 | FBgn0032497.5 - FBgn0032497.6 | GATATAGCAGCACCACATCC | GACGTGTGTAGACGTTCACC |
| CG6145 | CG6145.6 - CG6145.5 | CGTGTGTTGGATGATTCTCT | CTCGGTCGTTTACCTACCTT |
| CG6145 | CG6145.3 - CG6145.2 | ACCTGTTCCTGGAAGTTGTC | AGTTCCAAAAAGTGCATCAA |
| CG6767 | CG6767.9 - CG6767.3 | TTGGCTAGTTTCTTGCTTTC | CTCAGCCTCTCGTGTTACAG |
| CG6767 | CG6767.3 - CG6767.2A/B | TATGGAAAGCAAGGAATCAC | TGGTTACGAAGAAGTTCAGC |
| CG7378 | FBgn0030976.3 - FBgn0030976.4 | CACAGACAGATACAGCAGCA | GCGTCTTAGTCCTGGAAGAG |
| CG7378 | FBgn0030976.5 - FBgn0030976.6 | ACAGCTACTACCGGGATATG | GTAGAAGTAGCGCGAGATG |
| CG7852 | CG7852.1 - CG7852.2 | TACTCCGCGATATTTGTTTG | CCTCCGTTTCTCTGTCTCTC |
| CG7852 | CG7852.4 - CG7852.5 | ACGACATCAACAAGGACAAG | CTGAAGAATCAAAGCAATGG |
| dlg | FBgn0001624.9 - FBgn0001624.10 | GCAACAACACTACCACCAAC | ATTTTCCGTGTCCGATTC |
| dlg | FBgn0001624.2 - FBgn0001624.11 | AACTCAAGGAAAAGCCAAAG | CTCGAAATTGCGAACTAATG |
| dlg | FBgn0001624.1 - FBgn0001624.2 | AAACATAAAATGCCCGAAGG | TTTGCTTACTTCGCGTTCAC |
| dlg | FBgn0001624.1 - FBgn0001624.4 | TAGTGTTCCCGTGAGCAGAG | CCTTTTAGTTTCGCCAGCAC |
| dlg | FBgn0001624.7 - FBgn0001624.18 | TATTGGATGACTCGAAGAGC | GCTTGGTGATGTAGATGGAG |
| dlg | FBgn0001624.2 - FBgn0001624.6 | ACGCACACGTCAATGTAACC | AAAGTCTCGCGTGGTAGTCC |
| dlg | FBgn0001624.5 - FBgn0001624.6 | GCAGTTGAAGGCCAAGAGC | AAAGTCTCGCGTGGTAGTCC |
| dlg | FBgn0001624.6 - FBgn0001624.7 | TGCGTATTGCAATCGAACG | CCGCTGTCTTTTGTTGTATGC |
| dlg | FBgn0001624.7 - FBgn0001624.10 | CGAAGAGCATACAACAAA | ATTTTCCGTGTCCGATTC |
| dlg | FBgn0001624.16 - FBgn0001624.18 | TGACAACGAGGAAAAAGAAGC | ACCAGCTATCATCGCCATTC |

| | | | |
|---|---|---|---|
| dlg | FBgn0001624.18 - FBgn0001624.19 | GCGACAATGGCATCTATGTG | CGTTCTCCAGGTTCTTCTCG |
| dre4 | CG1828.16 - CG1828.7 | GGACTTTAAATGCGAAGAGG | TTGGCTCTGATGAGGAATC |
| dre4 | CG1828.6 - CG1828.5 | TCATGATCTTTTGCCAGTTC | CCATTCCTATGAACATGCTG |
| det | FBgn0024242.20 - FBgn0024242.21 | AGCACCGCACTGTTCAGCGG | GCTCGCTGCAGGACACATTGG |
| det | FBgn0024242.18 - FBgn0024242.20 | GCCGCAGCTCAACGAACTGG | GCTGAACAGTGCGGTGCTAGGG |
| det | FBgn0024242.34 - FBgn0024242.35 | CCGTGGGCGAGGCCAATGAG | ATGGCGGCTTCACGCTCTGG |
| Ect4 | FBgn0085402.9 - FBgn0085402.10 | CGTGAGCGTTTCAGCAAATC | TGTGTGCGGTGTTGTTTCTC |
| Ect4 | FBgn0085402.16 - FBgn0085402.18 | GAACGCAAAGTATGCAAAAC | CGTTTGTGTGACTGAATGTG |
| Ect4 | FBgn0085402.11 - FBgn0085402.12 | CATGGCCAAGACGCTAGATG | TCGATGAAAACGGAGAAACC |
| eIF4G | FBgn0023213.10 - FBgn0023213.11 | AGGAGCAACTATGGACTCAG | ATATTACCAACGCCCCAAAG |
| eIF4G | FBgn0023213.15 - FBgn0023213.17 | CTGTAGGAAATCGCCAAAC | TCGGATAGGCACATACAAAC |
| eIF4G | FBgn0023213.8 - FBgn0023213.9 | TGTTTCCAGATTTTGGTCAG | AGTCAACGTGGTGATAATGG |
| Eip75B | CG8127.3 - CG8127.7 | CTGTGGCTAGTTCGGTAAAG | CGATACTGGATCTTTTGCTG |
| Eip75B | CG8127.2 - CG8127.1 | GAAACCGGAGGCCTTATC | TTCTTCTTCGCACATCTTTC |
| Eip75B | CG8127.5 - CG8127.4 | AGTCAAACATGCAGATCAGG | AGTTCACCAAGGAGAAGGTC |
| ens/CG14998 | FBgn0035500.9 - FBgn0035500.10 | CAAGCGGTTCATCGAGTAG | GACCTTATGCCGACAATACC |
| ens/CG14998 | FBgn0035500.4 - FBgn0035500.5 | TCGTGGAGAAGCTATTGATG | CTTAAGCGGGAGGAGAAG |
| Fas3 | FBgn0000636.6 - FBgn0000636.8 | ATCAATCAAGACGCCCACA | TACGCTTCCCTCCACAATA |
| Fas3 | FBgn0000636.3 - FBgn0000636.4 | GCGAGGGCTACTTCAACG | GATACTGCGGCTCATACGC |
| Fim | CG8649.2 - CG8649.8 | AGTTCTCCTTCTTGGAGACG | TAATGGCCGATATTCTCAAC |
| Fim | CG8649.3 - CG8649.2 | AATCGGAGTAAAGCCAGTTC | CGTCAACGGAGTCTACAAAC |
| fus | CG8205.8 - CG8205.7 | TACTTGTGTTCGGAATGGAC | TCGTCTACTGGCCCTACC |
| fus | CG8205.4 - CG8205.3 | TATCCCCAGTGGTGAAGAAG | GAGGCTTTGATACGCTTTG |
| fz2 | FBgn0016797.3 - FBgn0016797.4 | TTGATTTTCTGGTTCATTCG | GTAAGCAGACGTGAAATTGG |

| | | | |
|---|---|---|---|
| fz2 | FBgn0016797.1 (internal primer set not exon boundary) | ATCAGGACCGAGTAGAGAGG | GTGTAGGAGCTGGTGTCAAG |
| gce | FBgn0030626.1 - FBgn0030626.3 | GCTGCTAAAATGCAAAGAAC | TTTACGCCTCTATCCGTCTC |
| gce | FBgn0030626.2 - FBgn0030626.4 | AGACGGATAGAGGCGTAAAC | TTGAGATCCTCAATGTCCTG |
| hts | CG9325.12 - CG9325.11 | GATGTGCTGTCGTATTGTTG | CACCAGGTCATTGAGATCC |
| hts | CG9325.6 - CG9325.5 | GACTGCAGCCGATATAGATG | TTCCTGGTTAACCCATACG |
| Ih | CG8585.1-A/B - CG8585.2-A/B | AGCGGTCCTGGAGCCTTCGG | TGCTGGTGCGAGTGCGAGTG |
| Ih | CG8585.10 - CG8585.11 | GGCCTACCGCAAGCTGCCAC | GGCACTGACGCCACGAGGG |
| jbug | CG30092.10B/A - CG30092.19 | CAGCTTTGACTTGCTAGGAG | GAATCGGATTTTGATTACCC |
| jbug | CG30092.5 - CG30092.4 | GCTCTCCCGCTTATAGTACC | ACTTCAATGGCTACGATGTC |
| jbug | CG30092.9 - CG30092.7 | GGCTATGAACCTCTGACCTC | TAGCGCCTATATCACTACGG |
| jbug | CG30092.17 - CG30092.11 | TTATTTTGTGGTTCGGACTG | ATTACGGTTAGCATCAGTGG |
| jbug | CG30092.19 - CG30092.18 | ATGGCTGGGTAATCAAAATC | GGAAAATTGGCAAGAAAAAC |
| jbug | CG30092.11 - CG30092.10A/B | GGCATTTTGTTGAATGACAC | ATGTGGAGTACAACGGATTC |
| l(1)G0232 | CG32697.9A/B - CG32697.8 | GCACCTAGATTGTTGCAGAG | TCGAAAGGTGATTTGAGTTG |
| l(1)G0232 | CG32697.5 - CG32697.4 | GCACTCAACGCTAATTGTTC | CTGACCAAGAATCGCTACAC |
| loco | CG5248.3 - CG5248.2 | AGATCCGAGTCCAGCATTAG | GAAGCTACGAACAGTGAACC |
| loco | CG5248.4 - CG5248.3 | GCAGACAAGTTTGAGTGAGG | TATCAGCGTTTCATTCGTTC |
| LpR2 | FBgn0051092.11 - FBgn0051092.12 | ACGGCAGTGTGATAATGTTC | GTAAAAGTCAAGGCGAGAGG |
| LpR2 | FBgn0051092.7 - FBgn0051092.9 | CTGCAGTCTGAATCTCCATC | TGTGTTCCTCCTTGATGTTC |
| LpR2 | FBgn0051092.4 - FBgn0051092.5 | TCACTGTACCCATCTTGAGG | GCACGTGTACCATCCTTATC |
| LpR2 | FBgn0051092.13 - FBgn0051092.14 | CACTGGAACTGCTTCTCATC | CATTAATTCGGTGCTAAACG |

| | | | |
|---|---|---|---|
| LpR2 | FBgn0051092.6 - FBgn0051092.7 | TAATTTCATCACTGCCATCC | GATTCAGACTGCAGAAATGG |
| ltd | CG8024.2 - CG8024.10/7 | AGTTCTGGCTGAAGAACTGG | GACCATGGATCAGGAAATG |
| ltd | CG8024.2 - CG8024.7 | AGTTCTGGCTGAAGAACTGG | CTATCAGTCGGTGGCTAGTG |
| ltd | CG8024.10 - CG8024.8 | TTCTCTGGCTCTGTGGTATC | AGGAAGAGGCTCAGGAAAC |
| ltd | CG8024.2 - CG8024.6 | GAAGAACTGGTGCACATAGC | ACATGGACATCAACAGGAAC |
| ltd | CG8024.2 - CG8024.5A/B | AGTTCTGGCTGAAGAACTGG | GACCGTTTAAGCCCTATGAC |
| ltd | CG8024.3 - CG8024.2 | GTTATGATGCCCTGCTTTTC | GCTATGTGCACCAGTTCTTC |
| Mctp | FBgn0034389.1 - FBgn0034389.2 | GTCCTGAATGGGAGAATAGG | CCATGTGGTGATATTTGGAG |
| Mctp | FBgn0034389.4 - FBgn0034389.5 | TCCTTATGTCAAGTTCAAGGTC | CATAAAATCGTCCTGCAGTC |
| Mctp | FBgn0034389.8 - FBgn0034389.9 | CGGATATAGGTGGAAAATCC | TTAGGAGGGGAATGACTAGC |
| Mctp | FBgn0034389.7 - FBgn0034389.8 | TATTGCTGGTTAAAGCCAAG | CTTCATCAAACAGGTGCAAG |
| Mical | CG33208.11 - CG33208.10 | CTGCTCCATGTTTTTAATCC | ATTTCTACAAGAGCCTTCAGC |
| Mical | CG33208.24 - CG33208.14 | TTACCATGTTGTTCATGTGG | AGGGTGAGGAATTCAAGAAG |
| Mical | CG33208.16 - CG33208.15 | GTTGTGGTTGGTACTTCTGG | TGAGATATCCACCGACTCTG |
| mth | CG6936.2 - CG6936.1 | ATCCATGATATGGTCGTGAG | TTGATACACCGATCGAGAAC |
| mth | CG6936.9-CG6936.8 | ACACTCCTCGTGGATTTAGG | GCTAGAAAACACGACGACAC |
| mth | CG6936.4 - CG6936.3 | TAGCAGATGAAGCATTTTCC | CCTTCAGCATCTTACATTCG |
| Nhe2 | CG9256.13 - CG9256.14 | TTATAGCCGACACGCAGTAG | CTGAAGTTCATTTGCATCTTG |
| Nhe2 | CG9256.2 - CG9256.3 | AAATGAAATCGATGCAGAAG | TCCAGATTCCAATGATGAAC |
| Pax | CG31794.37 - CG31794.8 | GTGCTGTGTTGCTCCTAGAC | GCAACGACTACTTCGAGATG |
| Pax | CG31794.8 - CG31794.35 | CGACTCATATTGGCCTGTAG | TGATTCAGTCGAGGCAAG |
| Pax | CG31794.3 - CG31794.2 | CATAACCTGGAGTGTTGTCC | ATATCATCGTCGTTCCACAG |
| Pax | CG31794.7 - CG31794.6 | TCAGATCATCCAGTTCCTTG | AAACGAAGACGGAGAAGTTG |

| | | | |
|---|---|---|---|
| PHGPx | CG12013.6 - CG12013.4 | TGAGTCTCAGGCAGTTTCAG | ACCTTGCCCTTGTACTTTTC |
| PHGPx | CG12013.4 - CG12013.2 | AGCTGACGGATCTAAAGGAG | GGGAATCAAAACCAGTAAGG |
| Pino | CG4710.3 - CG4710.2 | GACATATTGGCTGCTGATG | GCATCATTTTCGATGTGG |
| Pino | CG4710.5 - CG4710.4 | ACTTCCGGGTAGCTGATG | TCACAATCTCGGTCTGTCTG |
| pnt | CG17077.5 - CG17077.4 | TCATGGAGACCAATGAGAAC | TCCAAAAATGCAACATAACC |
| pnt | CG17077.4 - CG17077.3 | TCAGTACCTCGTTGACCTTG | AAGGATCTGTGCAGTTTGTC |
| pnt | CG17077.4 - CG17077.12 | AAGGATGCCTTCAGTACCTC | TCAATTGGCCTGACTATGAC |
| pnt | CG17077.7 - CG17077.11 | TATTTTCGTTGTTGCTTTGC | GATGCGAATGCCTACTACAC |
| pnt | CG17077.8 - CG17077.7 | GATCACAGACAAAGCGGTAG | CAGAGTTTCATTTCGTGGAC |
| ps | FBgn0026188.9 - FBgn0026188.10 | TCAGTTTGTTCTGGAATTGAGG | GTGCGACTTGGACATCTTGA |
| ps | FBgn0026188.10 - FBgn0026188.11 | GGGTCAAGATGTCCAAGTCG | AGGATACGTTCAGGCAGGTG |
| Pvr | CG8222.4 - CG8222.3 | GACCTTTCGAGCTGATTTTC | AAGTGACGGCTGACTATGTG |
| Pvr | CG8222.10 - CG8222.9 | CATCCTTGTCGTTGTGGTAG | TCATCAATGGAGAGGATTTG |
| Rbp9 | CG3151.1 - CG3151.3 | ACCGCTATCTTGGCTTTATC | GTCGTGTTCCATTGATGTTC |
| Rbp9 | CG3151.37 - CG3151.44 | AACCCATAACCGTCAAGTTC | CAGGTTGTAGACGAAAATGC |
| rols | CG32096.6 - CG32096.5 | TCTTCAGGTTGTGATTGCTC | GCGGTTCTTTAACCACAAG |
| rols | CG32096.8/10/14 - CG32096.7 | TCATTCAATGCACCAAGTTC | GCTATCAGTATGCCCTACGC |
| scb | CG8095.9 - CG8095.2 | CTGATCTCTCACATCGAAGC | AGGACATATGGACTGCACAC |
| scb | CG8095.6 - CG8095.7 | TTCGATGTGAGCCAACTAAG | GGCAATGGGTATGTAGAATG |
| shep | CG32423.4 - CG32423.3 | CGCGAATGTAGAGATTTGTC | GGCAGCCTACCGCTATAC |
| shep | CG32423.9 - CG32423.8 | TATCACCGCTAACCATCATC | AATGCAGCAGAATGGAGTC |
| shot | CG18076.34 - CG18076.27 | TCGTAGAGCTTTTCGTTGTC | TTTAGGGTTGACTTCAGTGC |
| shot | CG18076.9 - CG18076.28 | AAGGGCATTAATCTGGTTGC | CCGAAGAACAACATCGACCT |

| | | | |
|---|---|---|---|
| shot | CG18076.23 - CG18076.22 | GAAGCCGAAGCCTGTTAG | CACCATTACCCGCAAAAC |
| shot | CG18076.30 - CG18076.29 | CGGGTGGCACATTAAAAGAC | CACTTACGCGCAGTTAACGA |
| shot | CG18076.36 - CG18076.35 | AATGATGGTGGTACGGATG | ATCGGAAGGTTCAACTTTTG |
| shot | CG18076.11 - CG18076.10 | GTGCTGTCGGTCTTGTTG | CAAGTCGTTCAAGGATGAAG |
| sun | CG9032.4 - CG9032.2 | TAATTTGACGATGGAACACC | CATCCAATACTCCAACATCG |
| sun | CG9032.2 - CG9032.1 | CGATGTTGGAGTATTGGATG | ACTGACCGGATTCAGAAAG |
| Tm1 | CG4898.17 - CG4898.15-A | TCAAATCAAGCGAGCACATC | TGGACTTCTCCTTCGCCTTA |
| Tm1 | CG4898.2 - CG4898.3 | GACATGGATGCCATCAAGAA | CGGTCTGGATCTTCTTCTGC |
| Tm1 | CG4898.4 - CG4898.5 | GCATCCAGTTGCTCGAAGA | AGCCTCCTCAGCAAGGAAAC |
| Tm1 | CG4898.6 - CG4898.7 | CATGGTTGAAGCCGATTTG | CTTCCAGGGACTTCAGGTTG |
| Tm1 | CG4898.8 - CG4898.9 | CCAACGTGAGGAGGAGTACAA | AGCCTGTCGACTTCCTTCTG |
| Tm1 | CG4898.1 - CG4898.2 | CGTTCAAGTCGCGGATAACT | TCCTGCTCGCAGACGAGA |
| tmod | FBgn0082582.1 - FBgn0082582.3 | GACCCGTGATCGTGTAATTCGCCTC | TCCGCTGGGCTTTTGGCAGC |
| tmod | FBgn0082582.1 - FBgn0082582.4 | CCCGTGATCGTGTAATTCGCCTCG | TCGGCGGAGCCACCGAAAAC |
| tmod | FBgn0082582.2 - FBgn0082582.4 | GAAGGCTCCAGGCATCTGCG | GGGCATCGGGTTCGGGCTTG |
| tmod | FBgn0082582.6 - FBgn0082582.9 | ACGTGGAGGTGGAGGAGCAGG | GGGGGTGGTGAGCGTGGTTG |
| tmod | FBgn0082582.9 - FBgn0082582.10 | GCTCACCACCCCCGCCAAAC | TCCGGCTGATCCGGGGTCTC |
| Treh | CG9364.10 - CG9364.3 | ATATAGTCCGCGTTCACATC | CTTGAGATCCTCACTGCTTG |
| Treh | CG9364.11 - CG9364.12 | AGTGAAATGCGAAAACAAAC | AATTCACGCGGCTAATAAAG |
| Treh | CG9364.13 - CG9364.3 | CGTGAATTGTCCAGTGAAAG | TCAGCTTCATGTCCACAAAC |
| Treh | CG9364.12 - CG9364.3 | ACCACCGACTATAACAATGC | AATTGTTCAGCTTCATGTCC |
| Treh | CG9364.4 - CG9364.5 | GGCCTAAACAACCTGAACAC | TGATGATGCTGACAATGATG |
| trio | CG18214.9 - CG18214.22/23 | CAGCGGTTTGCTACTTGG | GGATAATGGAAATGCAAATG |
| trio | CG18214.4 - CG18214.3 | AGGGCTGAGATTGCTGTC | CATCGAGGACTTCTTCTGG |
| trol | FBgn0001402.1 - FBgn0001402.2 | AATTTTTGTGGATTTTCTTGG | GATACAAGCACAATTTGAACG |

| | | | |
|---|---|---|---|
| trol | FBgn0001402.20 - FBgn0001402.21 | AGCTAGGACCTTCGGTATTG | GGTCTTTCTTGTGAGCAGTG |
| Trp1 | CG4758.5 - CG4758.2 | GCAGGATATTTACGACGATG | TCGGTGAATTTGGACTTTAG |
| Trp1 | CG4758.2 - CG4758.3 | GGCTCTGACAATTGTACGTC | CGCTTTCTTTTTCCTTCTTC |
| ttk | CG1856.4 - CG1856.5 | AAAGAACTCCAAGGATCACC | GTTGCTGTTATTGCTTCCTG |
| ttk | CG1856.7 - CG1856.2 | TGGTGTTTCCAACGAATAAG | GAAGACGGACAGAAGGTTG |
| ttk | CG1856.8 - CG1856.2 | TGTAAAGCAGAGCGTATGTG | TTCGGATTCTTACTCCTTGG |
| ttk | CG1856.3 - CG1856.4 | AAATGTGCAACGAGTCCTAC | TTGTCTTCATGATGGCTCAC |
| tws | CG6235.3 - CG6235.2 | AATCTGTGAGAAGCACCAAG | AAGGACGTTTATGCGACAG |
| tws | CG6235.4 - CG6235.3 | GACCTTCCACAATTTGACTG | GATTAAAGGCGCCCTAGAC |
| Vha44 | CG8048.2 - CG8048.10 | TCCAGTATTCCGACATCATC | GAGTAGAGGAAAAGCCCAAC |
| Vha44 | CG8048.3 - CG8048.2 | CCACATAGGTGTCCAGTTTG | AGACGTACGACACGATGAAC |
| VhaSFD | CG17332.7 - CG17332.2 | AGTGTGGGACTGTACTGGTG | CCGACCTGAACTTCTATCTG |
| VhaSFD | CG17332.4 - CG17332.3 | ATAATGCGTGTCACCTTCTC | TGAGTATCGATTTGCCTTTG |
| vir-1 | FBgn0043841.1/2 - FBgn0043841.3 | ATTCTTCACGCTCCTTCTTC | TCAGTACGCGTAAAGAGGTG |
| vir-1 | FBgn0043841.2 - FBgn0043841.5 | TTGTGTTTGTTGGTGTTGTG | AGAAGTTCGAGGAGTTCGAC |
| vir-1 | FBgn0043841.4 - FBgn0043841.6 | AAAGACGCTGTTGCTATCAC | CTGTAATTGGCAGGGATTC |
| wls | CG6210.3 - CG6210.1 | ATTCTCATTCGACTGTGTGG | TAGGAATATTGGCGACAAGC |
| yki | FBgn0034970.5 - FBgn0034970.6 | ATGGTCGATGTGGTAGTTTG | ATCCAATATCGCCAGCAG |
| yki | FBgn0034970.4 - FBgn0034970.6 | CACTAAGTTATCTGGGCAGTC | ACTACAAAATCTACGCAGTGG |
| yki | FBgn0034970.6 - FBgn0034970.7 | TCCCACTGCGTAGATTTTG | AGACAGTGCACAAGAAGCAG |
| Zasp | FBgn0083919.10 - FBgn0083919.13 | ACGAGCATAGCCCTGGTTC | GCAACAGCATCATCATCAG |
| Zasp | FBgn0083919.12 - FBgn0083919.14 | CCGTGTTGTTGTGTTTCAG | AGCCTGTCATCCATTTCC |
| Zasp | FBgn0083919.14 - FBgn0083919.16 | ACTGGAAATGGATGACAGG | GGCTACTCCAATGGAAACTC |
| Zasp | FBgn0083919.15 - FBgn0083919.11 | CGGACACAGTTACGGACACTT | CAGTAGCAACGGTAGCACCA |
| Zasp | FBgn0083919.25 - FBgn0083919.24 | TGGACTTGGATCTCGGTTTC | TAAAAGAAGATGGCCCAACC |

| | | | |
|---|---|---|---|
| Zasp | FBgn0083919.26 - FBgn0083919.24 | TGGACTTGGATCTCGGTTTC | TAAAAGAAGATGGCCCAACC |
| Zasp | FBgn0083919.6 - FBgn0083919.4 | ATCTGCACATTGCAGCTGTT | TGGATTCGTTGAGGAGAAGG |
| Zasp | FBgn0083919.20 - FBgn0083919.21 | CCTGGTATTCCTTCTCGTTC | TGAATCTATTGCGGAAACAC |
| zfh1 | CG1322.4 - CG1322.5 | GCATCAAAGTAGAGCAAGTGTTG | GGTTGTGGTGGAGCAAAGTG |
| zfh1 | CG1322.5 - CG1322.6 | GATGCCTTCCTGGTCAAGTG | ATGGACTGGCTGCTGGTG |
| zfh1 | CG1322.2 - CG1322.8 | GGAGACTTCGCTTCCAACC | GCCTTCGGACACTCTATGC |
| zip | CG15792.7 - CG15792.5 | ACTTCCACCATCTTCAATCC | AGGTCATCCAATTTCTAGCC |
| zip | CG15792.8 - CG15792.5 | TATCGTTTTTGACCGTCTTG | CCAATTTCTAGCCTATGTGG |
| zip | CG15792.4 - CG15792.3 | ATCGGAAGTTTCTTGTACGG | AGATGAATCCGCCTAAGTTC |

# References

Anders, Simon, Alejandro Reyes, and Wolfgang Huber. 2012. "Detecting Differential Usage of Exons from RNA-Seq Data." *Genome research* 22(10): 2008–17.

Andrews, Justen, and Brian Oliver. 2002. "Sex Determination Signals Control Ovo-B Transcription in Drosophila Melanogaster Germ Cells." *Genetics* 160(2): 537–45.

Aravin, Alexei A., Alexei V. Tulin Natalia M. Naumova, and Yakov M. Rozovsky and Vladimir A. Gvozdev Vasili V. Vagin. 2001. "Double-Stranded RNA-Mediated Silencing of Genomic Tandem Repeats and Transposable Elements in the D. Melanogaster Germline Alexei A. Aravin* † , Natalia M. Naumova* † , Alexei V. Tulin* ‡ , Vasilii V. Vagin* † , Yakov M. Rozovsky* and Vladimir A. Gvozdev*." *Current Biology* 11(13): 1017–27.

Bate, M., and Martinez Arias, A., (1993) "The Development of Drosophila melanogaster." *Cold Spring Harbor Laboratory Press*, Cold Spring Harbor, NY.

Bilder, D, M Li, and N Perrimon. 2000. "Cooperative Regulation of Cell Polarity and Growth by Drosophila Tumor Suppressors." *Science* 289(5476): 113–16.

Bloor, J W, and D P Kiehart. 2001. "Zipper Nonmuscle Myosin-II Functions Downstream of PS2 Integrin in *Drosophila* Myogenesis and Is Necessary for Myofibril Formation." *Developmental biology* 239: 215–28.

Bopp, D et al. 1999. "Recombination and Disjunction in Female Germ Cells of Drosophila Depend on the Germline Activity of the Gene Sex-Lethal." *Development* 126(24): 5785–94.

Borghese, Lodovica et al. 2006. "Systematic Analysis of the Transcriptional Switch Inducing Migration of Border Cells." *Developmental Cell* 10(4): 497–508.

Bossinger, O et al. 2001. "Zonula Adherens Formation in Caenorhabditis Elegans Requires Dlg-1, the Homologue of the Drosophila Gene Discs Large." *Developmental biology* 230(1): 29–42.

Bozzetti, M P et al. 1995. "The Ste Locus, a Component of the Parasitic Cry-Ste System of Drosophila Melanogaster, Encodes a Protein That Forms Crystals in Primary Spermatocytes and Mimics Properties of the Beta Subunit of Casein Kinase 2." *Proceedings of the National Academy of Sciences of the United States of America* 92(13): 6067–71.

Breitwieser, Wolfgang, Finn Hugo Markussen, Heinz Horstmann, and Anne Ephrussi. 1996. "Oskar Protein Interaction with Vasa Represents an Essential

Step in Polar Granule Assembly." *Genes and Development* 10(17): 2179–88.

Camara, Nicole, Cale Whitworth, and Mark Van Doren. 2008. " The Creation of Sexual Dimorphism in the Drosophila Soma." *Current Topics in Developmental Biology* 83(08): 65-107.

Casper, Abbie L, and Mark Van Doren. 2009. "The Establishment of Sexual Identity in the Drosophila Germline." *Development* 136(22): 3821–30.

Chau, J., L. S. Kulnane, and H. K. Salz. 2009. "Sex-Lethal Facilitates the Transition From Germline Stem Cell to Committed Daughter Cell in the Drosophila Ovary." *Genetics* 182(1): 121–32.

Chau, Johnnie, Laura Shapiro Kulnane, and Helen K Salz. 2012. "Sex-Lethal Enables Germline Stem Cell Differentiation by down-Regulating Nanos Protein Levels during Drosophila Oogenesis." *Proceedings of the National Academy of Sciences of the United States of America* 109(24): 9465–70.

Cheng, Jun et al. 2008. "Centrosome Misorientation Reduces Stem Cell Division during Ageing." *Nature* 456(7222): 599–604.

de Cuevas, M, and A C Spradling. 1998. "Morphogenesis of the Drosophila Fusome and Its Implications for Oocyte Specification." *Development* 125(15): 2781–89.

Dobens, Leonard L, and Laurel A Raftery. 2000. "Integration of Epithelial Patterning and Morphogenesis in Drosophila Ovarian Follicle Cells." *Developmental Dynamics* 218: 80–93.

Erickson, James W, and Jerome J Quintero. 2007. "Indirect Effects of Ploidy Suggest X Chromosome Dose, Not the X:A Ratio, Signals Sex in Drosophila." *PLoS biology* 5(12): e332.

Eun, Suk Ho et al. 2013. "MicroRNAs Downregulate Bag of Marbles to Ensure Proper Terminal Differentiation in the Drosophila Male Germline." *Development* 140(1): 23–30.

Firestein, B L, and C Rongo. 2001. "DLG-1 is a MAGUK Similar to SAP97 and Is Required for Adherens Junction Formation." *Molecular biology of the cell* 12(11): 3465–75.

Fuller, M T. 1998. "Genetic Control of Cell Proliferation and Differentiation in Drosophila Spermatogenesis." *Seminars in cell & developmental biology* 9(4): 433–44.

Gan, Qiang et al. 2010. "Dynamic Regulation of Alternative Splicing and Chromatin Structure in Drosophila Gonads Revealed by RNA-Seq." *Cell research* 20(7): 1–21.

Gilboa, Lilach, and Ruth Lehmann. 2006. "Soma–germline Interactions

Coordinate Homeostasis and Growth in the Drosophila Gonad." *Nature* 443(7107): 97–100.

Gönczy, P, E Matunis, and S DiNardo. 1997. "Bag-of-Marbles and Benign Gonial Cell Neoplasm Act in the Germline to Restrict Proliferation during Drosophila Spermatogenesis." *Development* 124(21): 4361–71.

Harris, Robin E. et al. 2011. "Brat Promotes Stem Cell Differentiation via Control of a Bistable Switch That Restricts BMP Signaling." *Developmental Cell* 20(1): 72–83.

Harris, Tony J C, and Mark Peifer. 2004. "Adherens Junction-Dependent and -Independent Steps in the Establishment of Epithelial Cell Polarity in Drosophila." *The Journal of cell biology* 167(1): 135–47.

Hashiyama, Kazuya, Yoshiki Hayashi, and Satoru Kobayashi. 2011. "Drosophila Sex Lethal Gene Initiates Female Development in Germline Progenitors." *Science* 333(6044): 885–88.

Hiller, Mark et al. 2004. "Testis-Specific TAF Homologs Collaborate to Control a Tissue-Specific Transcription Program." *Development* 131(21): 5297–5308.

Hinson, S, and R N Nagoshi. 1999. "Regulatory and Functional Interactions between the Somatic Sex Regulatory Gene Transformer and the Germline Genes Ovo and Ovarian Tumor." *Development* 126(5): 861–71.

Hough, C. D., D. F. Woods, S. Park, and P. J. Bryant. 1997. "Organizing a Functional Junctional Complex Requires Specific Domains of the Drosophila MAGUK Discs Large." *Genes & Development* 11(23): 3242–53.

Humbert, P O et al. 2008. "Control of Tumourigenesis by the Scribble/Dlg/Lgl Polarity Module." *Oncogene* 27(55): 6888–6907.

Humbert, Patrick, Sarah Russell, and Helena Richardson. 2003. "Dlg, Scribble and Lgl in Cell Polarity, Cell Proliferation and Cancer." *BioEssays : news and reviews in molecular, cellular and developmental biology* 25(6): 542–53.

Huynh, Jean-René, and Daniel St Johnston. 2004. "The Origin of Asymmetry: Early Polarisation of the Drosophila Germline Cyst and Oocyte." *Current biology : CB* 14(11): R438–49.

Inaba, Mayu et al. 2010. "E-Cadherin Is Required for Centrosome and Spindle Orientation in Drosophila Male Germline Stem Cells." *PloS one* 5(8): e12473.

Janic, A et al. 1994. "Ectopic Expression of Germline Genes Drives Malignant Brain Tumor Growth in Drosophila." *Science* 79: 1165–73.

Jemc, Jennifer C. 2011. "Somatic Gonadal Cells: The Supporting Cast for the Germline." *Genesis 49:10: 753-775*.

Kai, Toshie, Dianne Williams, and Allan C Spradling. 2005. "The Expression

Profile of Purified Drosophila Germline Stem Cells." *Developmental biology* 283(2): 486–502.

Kelley, R.L., J. Wangf, L. Bellt, and M.I. Kurodat. 1997. "Sex Lethal Controls Dosage Compensation in Drosophila by a Non-Splicing Mechanism." *gene expression* 5: 2048–59.

Keyes, L N, T W Cline, and P Schedl. 1992. "The Primary Sex Determination Signal of Drosophila Acts at the Level of Transcription." *Cell* 68(5): 933–43.

Kim, Daehwan et al. 2013. "TopHat2 : Accurate Alignment of Transcriptomes in the Presence of Insertions , Deletions and Gene Fusions." *Genome Biology 14:R36*: 1–13.

King, R C et al. 1986. "Complementation between Alleles at the Ovarian Tumor Locus of Drosophila Melanogaster." *Developmental genetics* 7(1): 1–20.

Laprise, Patrick, Alain Viel, and Nathalie Rivard. 2004. "Human Homolog of Disc-Large Is Required for Adherens Junction Assembly and Differentiation of Human Intestinal Epithelial Cells." *The Journal of biological chemistry* 279(11): 10157–66.

Leatherman, Judith L, and Stephen Dinardo. 2010. "Germline Self-Renewal Requires Cyst Stem Cells and Stat Regulates Niche Adhesion in Drosophila Testes." *Nature cell biology* 12(8): 806–11.

Li, Y. et al. 2012. "Mei-P26 Regulates the Maintenance of Ovarian Germline Stem Cells by Promoting BMP Signaling." *Development* 139(9): 1547–56.

Lin, H, L Yue, and a C Spradling. 1994. "The Drosophila Fusome, a Germline-Specific Organelle, Contains Membrane Skeletal Proteins and Functions in Cyst Formation." *Development* 120(4): 947–56.

Lin, Haifan, and Allan C. Spradling. 1993. "GSC Division and Egg Chamber Development in Transplanted Drosophila Germaria." *Developmental Biology* 159: 140–52.

Lin, Ting-yi et al. 1996. "Coordinate Developmental Control of the Meiotic Cell Cycle and Spermatid Differentiation in Drosophila Males." *Development* 122: 1331–41.

Liu, Y, and D J Montell. 1999. "Identification of Mutations That Cause Cell Migration Defects in Mosaic Clones." *Development* 126(9): 1869–78.

Mahowald, a P. 2001. "Assembly of the Drosophila Germ Plasm." *International review of cytology* 203: 187–213.

Marsh, JL, and E. Wieschaus. 1978. "Is Sex Determination in Germ Line and Soma Controlled by Separate Genetic Mechanisms?" Nature 272.

McIntyre, L.M. et al. 2006. "Sex-Specific Expression of Alternative Transcripts in

Drosophila." *Genome Biology* 7(8): R79.

McKearin, D M, and A C Spradling. 1990. "Bag-of-Marbles: A Drosophila Gene Required to Initiate Both Male and Female Gametogenesis." *Genes & Development* 4(12b): 2242–51.

McKearin, D, and B Ohlstein. 1995. "A Role for the Drosophila Bag-of-Marbles Protein in the Differentiation of Cystoblasts from Germline Stem Cells." *Development* 121(9): 2937–47.

Mummery-Widmer, Jennifer L et al. 2009. "Genome-Wide Analysis of Notch Signalling in Drosophila by Transgenic RNAi." *Nature* 458(7241): 987–92.

Nagoshi, R N, J S Patton, E Bae, and P K Geyer. 1995. "The Somatic Sex Determines the Requirement for Ovarian Tumor Gene Activity in the Proliferation of the Drosophila Germline." *Development* 121(2): 579–87.

Ni, Jian-Quan et al. 2011. "A Genome-Scale shRNA Resource for Transgenic RNAi in Drosophila." *Nature methods* 8(5): 405–7.

Oliver, B, Y J Kim, and B S Baker. 1993. "Sex-Lethal, Master and Slave: A Hierarchy of Germ-Line Sex Determination in Drosophila." *Development* 119(3): 897–908.

Oliver, B, D Pauli, and a P Mahowald. 1990. "Genetic Evidence That the Ovo

Locus Is Involved in Drosophila Germ Line Sex Determination." *Genetics* 125(3): 535–50.

Papagiannouli, Fani, and Bernard M Mechler. 2009. "Discs Large Regulates Somatic Cyst Cell Survival and Expansion in Drosophila Testis." *Cell research* 19(10): 1139–49.

Patil, Veena S., and Toshie Kai. 2010. "Repression of Retroelements in Drosophila Germline via piRNA Pathway by the Tudor Domain Protein Tejas." *Current Biology* 20(8): 724–30.

Pek, J. W., a. Anand, and T. Kai. 2012. "Tudor Domain Proteins in Development." *Development* 139(13): 2255–66.

Raff, J W, W G Whitfield, and D M Glover. 1990. "Two Distinct Mechanisms Localise Cyclin B Transcripts in Syncytial Drosophila Embryos." *Development* 110(4): 1249–61.

Ren, Xingjie et al. 2013. "Optimized Gene Editing Technology for Drosophila Melanogaster Using Germ Line-Specific Cas9." *Proceedings of the National Academy of Sciences of the United States of America* 110(47): 19012–17.

Robinson, D N, K Cant, and L Cooley. 1994. "Morphogenesis of Drosophila Ovarian Ring Canals." *Development* 120(7): 2015–25.

Salz, H.K., T.W. Cline, and Paul Schedl. 1987. "Functional Changes Associated with Structural Alterations Induced by Mobilization of a P Element Inserted in the Sex-Lethal Gene of Drosophila." *Genetics* 117(2): 221.

Schupbach, T. 1982. "Autosomal Mutations That Interfere with Sex Determination in Somatic Cells of Drosophila Have No Direct Effect on the Germline." *Developmental Biology* 89(1): 117–27.

Schupbach, T. 1985. "Normal Female Germ Cell Differentiation Requires the Female X Chromosome to Autosome Ratio and Expression of Sex-Lethal in Drosophila Melanogaster." *Genetics*: 529–48.

Shapiro-Kulnane, L., a. E. Smolko, and H. K. Salz. 2015. "Maintenance of Drosophila Germline Stem Cell Sexual Identity in Oogenesis and Tumorigenesis." *Development* 142(6): 1073–82.

Siegrist, Sarah E, and Chris Q Doe. 2005. "Microtubule-Induced Pins/Galphai Cortical Polarity in Drosophila Neuroblasts." *Cell* 123(7): 1323–35.

Smendziuk, C. M., A. Messenberg, A. W. Vogl, and G. Tanentzapf. 2015. "Bi-Directional Gap Junction-Mediated Soma-Germline Communication Is Essential for Spermatogenesis." *Development* 142(15): 2598–2609.

Smith, James M. et al. 2004. "Expression of the Tudor-Related Gene Tdrd5 during Development of the Male Germline in Mice." *Gene Expression Patterns* 4(6):

701–5.

Song, Xiaoqing et al. 2004. “Bmp Signals from Niche Cells Directly Repress Transcription of a Differentiation-Promoting Gene, Bag of Marbles, in Germline Stem Cells in the Drosophila Ovary.” *Development* 131(6): 1353–64.

Song, Xiaoqing, Chun-Hong Zhu, Chuong Doan, and Ting Xie. 2002. “Germline Stem Cells Anchored by Adherens Junctions in the Drosophila Ovary Niches.” *Science* 296(5574): 1855–57.

Staab, S, A Heller, and M Steinmann-Zwicky. 1996. “Somatic Sex-Determining Signals Act on XX Germ Cells in Drosophila Embryos.” *Development* 122(12): 4065–71.

Steinmann-Zwicky, M, H Schmid, and R Nöthiger. 1989. “Cell-Autonomous and Inductive Signals Can Determine the Sex of the Germ Line of Drosophila by Regulating the Gene Sxl.” *Cell* 57(1): 157–66.

Storto, Patrick Daniel, and Robert Charles King. 1987. “Fertile heteroallelic combinations of mutant alleles of the otu Locus of Drosophila Melanogaster.” *Developmental Biology*: 210–21.

Tanentzapf, Guy, and Ulrich Tepass. 2003. “Interactions between the Crumbs, Lethal Giant Larvae and Bazooka Pathways in Epithelial Polarization.” *Nature cell biology* 5(1): 46–52.

Tazuke, Salli I et al. 2002. "A Germline-Specific Gap Junction Protein Required for Survival of Differentiating Early Germ Cells." *Development* 129(10): 2529–39.

Telonis-Scott, Marina et al. 2009. "Sex-Specific Splicing in Drosophila: Widespread Occurrence, Tissue Specificity and Evolutionary Conservation." *Genetics* 181(2): 421–34.

Trapnell, Cole et al. 2010. "Transcript Assembly and Quantification by RNA-Seq Reveals Unannotated Transcripts and Isoform Switching during Cell Differentiation." *Nature biotechnology* 28(5): 511–15.

Trapnell, Cole et al. 2012. "Differential Gene and Transcript Expression Analysis of RNA-Seq Experiments with TopHat and Cufflinks." *Nature protocols* 7(3): 562–78.

Van Doren, M, A L Williamson, and R Lehmann. 1998. "Regulation of Zygotic Gene Expression in Drosophila Primordial Germ Cells." *Current biology* 8(4): 243–46.

Venken, Koen J T et al. 2009. "Versatile P[acman] BAC Libraries for Transgenesis Studies in Drosophila Melanogaster." *Nature Methods* 6(6): 431–34.

Voronina, Ekaterina, Geraldine Seydoux, Paolo Sassone-Corsi, and Ippei Nagamori. 2011. "RNA Granules in Germ Cells." *Cold Spring Harbor*

*Perspectives in Biology* 3(12).

Wang, Eric T et al. 2008. "Alternative Isoform Regulation in Human Tissue Transcriptomes." *Nature* 456(7221): 470–76.

Wawersik, Matthew et al. 2005. "Somatic Control of Germline Sexual Development Is Mediated by the JAK/STAT Pathway." *Nature* 436(7050): 563–67.

White-Cooper, H, M a Schäfer, L S Alphey, and M T Fuller. 1998. "Transcriptional and Post-Transcriptional Control Mechanisms Coordinate the Onset of Spermatid Differentiation with Meiosis I in Drosophila." *Development* 125(1): 125–34.

White-cooper, Helen. 2004. "Analysis of Meiosis and Morphogenesis." *Methods in Molecular Biology* 247.

Woods, D F et al. 1996. "Dlg Protein Is Required for Junction Structure, Cell Polarity, and Proliferation Control in Drosophila Epithelia." *The Journal of cell biology* 134(6): 1469–82.

Woods, D F, and P J Bryant. 1991. "The Discs-Large Tumor Suppressor Gene of Drosophila Encodes a Guanylate Kinase Homolog Localized at Septate Junctions." *Cell* 66(3): 451–64.

Woods, D F, J W Wu, and P J Bryant. 1997. "Localization of Proteins to the Apico-Lateral Junctions of Drosophila Epithelia." *Developmental genetics* 20(2): 111–18.

Xia, Laixin et al. 2010. "The Fused/smurf Complex Controls the Fate of Drosophila Germline Stem Cells by Generating a Gradient Bmp Response." *Cell* 143(6): 978–90.

Xie, Ting, and Allan C Spradling. 1998. " Decapentaplegic is Essential for the Maintenance and Division of Germline Stem Cells in the Drosophila Ovary." *Cell* 94(2): 251–60.

Yabuta, Yukihiro et al. 2011. "TDRD5 Is Required for Retrotransposon Silencing, Chromatoid Body Assembly, and Spermiogenesis in Mice." *Journal of Cell Biology* 192(5): 781–95.

Yacobi-Sharon, Keren, Yuval Namdar, and Eli Arama. 2013. "Alternative Germ Cell Death Pathway in Drosophila Involves HtrA2/Omi, Lysosomes, and a Caspase-9 Counterpart." *Developmental Cell* 25(1): 29–42.

Yamashita, Yukiko M, D Leanne Jones, and Margaret T Fuller. 2003. "Orientation of Asymmetric Stem Cell Division by the APC Tumor Suppressor and Centrosome." *Science* 301(5639): 1547–50.

Yang, Shu Yuan, Ellen M. Baxter, and Mark Van Doren. 2012. "Phf7 Controls

Male Sex Determination in the Drosophila Germline." *Developmental Cell* 22(5): 1041–51.

Yue, Lin, and Allan C. Spradling. 1992. "Hu-Li Tai Shao, a Gene Required for Ring Canal Formation during Drosophila Oogenesis, Encodes a Homolog of Adducin." *Genes and Development* 6: 2443–54.

# Shekerah Primus

Johns Hopkins University ▪ 3400 N Charles St ▪ Mudd Hall rm 302 ▪ Baltimore, MD 21218

973-699-2112 ▪ sprimus1@jhu.edu

## Education

Johns Hopkins University, Baltimore, MD
**Ph.D., Biology**
*Defended: December 2015*
Dissertation: "Investigating the role of *Sex lethal* in *Drosophila* germline sexual identity"

Montclair State University, Montclair, NJ
Bachelor of Science, Molecular Biology
Bachelor of Science, Science Informatics
*Graduated Summa Cum Laude: May 2008*

## Research Experience

Johns Hopkins University: Cell, Molecular, Developmental Biology, and Biophysics graduate program
***Graduate Student:*** *September 2009 – December2015*
Principal Investigator: Mark Van Doren

- Investigated the role of the gene, *discs large*, in the *Drosophila* germline using RNAi, phenotypic analysis using fluorescence microscopy, and by evaluating epistatic interactions
- Conducted RNA-seq to identify germline targets of the RNA-binding protein Sex lethal
- Determined how *Sex lethal* promotes female identity in the germline, partly through the analysis of transcriptome data
- Functional characterization of a new *Sex lethal* target in the germline, by analyzing CRISPR-Cas9 generated mutants, as well as overexpression and genetic interaction studies

University of Medicine and Dentistry of New Jersey, Newark, NJ
***Independent Undergraduate Research:*** *January 2008 - October 2008*
Principal Investigator: Nikhat Parveen

- Standardized a high-throughput fluorometric assay to test novel antimicrobials against Borrelia *burgdorferi*
- Identified antimicrobial agents effective in inhibiting the growth of *B. burgdorferi*

## Publications

---

**Research Papers:**

1. **Primus S**, Baxter K, Van Doren M. *Tdr5p* promotes male sexual identity in the *Drosophila* germline and is repressed in females by SXL. *(manuscript in preparation).*

2. **Primus S**, and Van Doren M. *Discs large* is needed for proliferation and survival of male germ cells in Drosophila *melanogaster*. *(manuscript in preparation).*

3. Leung W. et al. (2015) Drosophila muller f elements maintain a distinct set of genomic properties over 40 million years of evolution. G3 (Bethesda), 5(5): 719-40.

4. Cornell KA, **Primus S**, Martinez JA, Parveen N, (2009) Assessment of methylthioadenosine/S-adenosylhomocysteine nucleosidases of Borrelia burgdorferi as targets for novel antimicrobials using a novel high-throughput method. J. Antimicrob Chemother, 63(6): 1163-72.

**Abstracts:**

Conference Talks

1. **Shekerah Primus*** & Mark Van Doren. (2014). Investigating the role of *Sex lethal* in the germline of *Drosophila melanogaster*. Invited speaker* 55th Annual *Drosophila* Research Conference.

Conference Posters

2. **Shekerah Primus** & Mark Van Doren (2012). The significance of alternative splicing for germline sexual development. Presented at the Cold Spring Harbor Laboratory Germ Cells Meeting.

3. **Shekerah Primus** & Mark Van Doren (2011). The significance of alternative splicing to germline sex determination. Presented at the 52nd Annual Drosophila Research Conference.

## Teaching Experience

Instructor, JHU, Introduction to Laboratory Research, Lecture and Lab
Summer 2016
Instructor, JHU, Introduction to Laboratory Research, Lecture and Lab
Summer 2015
Teacher's Assistant, JHU, Developmental Biology, Lab
Spring 2011
Teacher's Assistant, JHU, Genetics, Lecture
Fall 2010
Teacher's Assistant, JHU, Genetics, Lab
Fall 2010